HAL
open science

Stable model reduction for linear variational inequalities with parameter-dependent constraints

Idrissa Niakh, Guillaume Drouet, Virginie Ehrlacher, Alexandre Ern

To cite this version:

Idrissa Niakh, Guillaume Drouet, Virginie Ehrlacher, Alexandre Ern. Stable model reduction for linear variational inequalities with parameter-dependent constraints. ESAIM: Mathematical Modelling and Numerical Analysis, 2023, 57 (1), pp.167–189. 10.1051/m2an/2022077 . hal-03611982v3

HAL Id: hal-03611982

https://hal.science/hal-03611982v3

Submitted on 12 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stable model reduction for linear variational inequalities with parameter-dependent constraints

Idrissa Niakh *†‡, Guillaume Drouet *, Virginie Ehrlacher †‡, Alexandre Ern †‡

September 9, 2022

Abstract

We consider model reduction for linear variational inequalities with parameter-dependent constraints. We study the stability of the reduced problem in the context of a dualized formulation of the constraints using Lagrange multipliers. Our main result is an algorithm that guarantees inf-sup stability of the reduced problem. The algorithm is computationally effective since it can be performed in the offline phase even for parameter-dependent constraints. Moreover, we also propose a modification of the Cone Projected Greedy algorithm so as to avoid ill-conditioning issues when manipulating the reduced dual basis. Our results are illustrated numerically on the frictionless Hertz contact problem between two half-spheres with parameter-dependent radius and on the membrane obstacle problem with parameter-dependent obstacle geometry.

Keywords — model reduction, variational inequalities, reduced basis method, contact problem, obstacle problem, inf-sup condition.

1 Introduction

Model reduction is a method for reducing computational costs by approximating a high-fidelity (HF) parameter-dependent model that is computationally expensive by another model that is much cheaper to solve; see [32, 33, 30, 38, 3, 12] and the more recent textbooks [25, 34]. The principle is to organize the calculations in a first offline phase which is a learning phase where expensive calculations are carried out on the HF model for a small number of values of the parameter drawn from a training set. The output of the offline phase is used to build the reduced-order model. Then, in the online phase, a large number of new instances of the parameter are considered for which the reduced model is solved instead of the HF model. Such studies are often carried out, for example, in the context of model calibration, where the aim is to find the optimal value for certain parameters so that the results of numerical simulations are as close as possible to those of the experiments.

The work developed herein is concerned with model reduction for linear parameter-dependent variational inequalities [22, 20, 40]. These inequalities are, for instance, encountered in several problems

*EDF R&D, 7 Boulevard Gaspard Monge, 91120 Palaiseau, France
†CERMICS, École des Ponts, 6-8 avenue Blaise Pascal, 77455 Marne-la-Vallée cedex 2, France
‡INRIA Paris, 2 Rue Simone Iff, 75012 Paris, France

in computational mechanics [27]. As example of applications, we will study the frictionless Hertz contact problem [24, 39] and the membrane obstacle problem [17, 18, 26, 31, 21]. We are particularly interested in the case where the problem is formulated by introducing Lagrange multipliers. This formulation leads to a so-called primal-dual strategy, whereby reduced bases are created for both the primal and the dual variables.

A first possibility in the primal-dual context is to generate the reduced primal and dual bases in a decorrelated manner. This is for instance the case in [7] where, after sampling pairs of primal and dual solutions for parameter values taken from a training set, the Proper Orthogonal Decomposition (POD) [13, 28] is used to compress the primal basis and a Non-negative Matrix Factorization (NMF) [29] is used to compress the dual basis. Let us also mention [16] where a hypereduction method for contact problems is proposed. Instead of the NMF, one can also consider the Angle Greedy [22, 11] and the Cone Projected Greedy (CPG) [9] algorithms for the compression of the dual basis. However, whatever the compression technique, if the primal and dual reduced bases are generated in a decorrelated way, the stability of the reduced problem is not guaranteed a priori. Specifically, it cannot be guaranteed that the resulting reduced bases are such that the operator associated with the constraint satisfies an inf-sup condition for the reduced primal and dual spaces.

In order to satisfy this inf-sup condition, a strategy for constructing the reduced primal basis according to the reduced dual basis has been proposed in [37, 19, 22, 36] in various contexts. The idea is to complete the reduced primal basis with as many functions as there are in the reduced dual basis, each of these functions being determined by a maximization problem in order to control the corresponding element of the reduced dual basis. These functions are often called supremizers; see, e.g., [8, 2]. In the case where the constraints are parameter-independent (which is the case in the above references), the completion of the reduced primal basis can be calculated only once in the offline phase. However, in the case where the constraints are parameter-dependent, the supremizers become parameter-dependent so that the reduced primal space has to be constructed during the online phase. This considerably reduces the efficiency of the online phase. An example of this situation is the frictionless contact problem between Hertz spheres, even under the assumption of small deformations, when the radius of one of the spheres depends on the parameter.

In this paper, we propose a strategy to approximate the parameter-dependent primal space ensuring inf-sup stability by a parameter-independent space. This latter space can thus be constructed only once in the offline phase. In addition, we establish a criterion to be checked during the construction in the offline phase so as to guarantee uniform inf-sup stability condition for the reduced problem (at least for all the parameters in the training set). We notice that an algorithm to build supremizers in the context of parameter-dependent constraints has also been devised in [14]; we discuss below the differences between this contribution and the present one.

A second, distinct contribution of the present work is a modified version of the CPG algorithm to build a cone with a wider aperture. The motivation is to achieve better approximation properties compared to the ones provided by the plain CPG algorithm. In our numerical experiments, this modification turned out to be instrumental in order to avoid ill-conditioning issues when manipulating the reduced dual basis. Another algorithm to enhance the aperture of a given cone has been recently devised in [6]; we discuss below the differences between this contribution and the present one.

The rest of this paper is organized as follows. In Section 2, we introduce the mathematical framework for parameter-dependent variational inequalities. Then, we formulate the reduced model in the standard cases where the reduced primal basis is decorrelated from the reduced dual basis or enriched by parameter-dependent supremizers to be computed online. In Section 3, we present our main result, namely the strategy to achieve stability by enriching the primal basis in the offline phase. In Section 4, we describe the new version of the CPG algorithm to enhance the aperture of the cone associated with the reduced dual basis. In Section 5, numerical results illustrating the efficiency of our approach are reported for the frictionless Hertz contact problem between two half-spheres with

parameter-dependent radius and the membrane obstacle problem with parameter-dependent obstacle geometry. In future work, we plan to extend our work to the case of nonlinear problems such as the contact problem under large deformations and the contact problem with Coulomb friction.

2 Setting

2.1 High-fidelity model

Let $\mathcal{V}$ and $\mathcal{W}$ be two finite-dimensional high-fidelity (HF) subspaces typically resulting from the finite element discretization of some Hilbert spaces. We denote by $\langle \cdot, \cdot \rangle_{\mathcal{V}}$ (resp. $\langle \cdot, \cdot \rangle_{\mathcal{W}}$) the inner product equipping $\mathcal{V}$ (resp. $\mathcal{W}$) and inducing the norm $\| \cdot \|_{\mathcal{V}}$ (resp. $\| \cdot \|_{\mathcal{W}}$). We denote by $\mathcal{V}'$ the dual space of $\mathcal{V}$ equipped with the duality product $\langle \cdot, \cdot \rangle_{\mathcal{V}',\mathcal{V}}$, and by $\| \cdot \|_{\mathcal{V}'}$ the associated norm. Let $\mathcal{W}^+ \subset \mathcal{W}$ be a positive cone containing 0. Let $\mathcal{D} \subset \mathbb{R}^m, m \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}$, be the parameter domain. We consider two uniformly bounded bilinear forms $a(\mu; \cdot, \cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ and $b(\mu; \cdot, \cdot) : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$ for all $\mu \in \mathcal{D}$, and two uniformly bounded linear forms $f(\mu; \cdot) : \mathcal{V} \to \mathbb{R}$ and $g(\mu; \cdot) : \mathcal{W} \to \mathbb{R}$ for all $\mu \in \mathcal{D}$. We assume that $a(\mu; \cdot, \cdot)$ is symmetric and uniformly coercive for all $\mu \in \mathcal{D}$. We consider the following well-posed constrained minimization problem: For all $\mu \in \mathcal{D}$, find $u(\mu) \in \mathcal{V}$ such that

$$u(\mu) = \underset{v \in \mathcal{K}(\mu)}{\operatorname{argmin}} \ \frac{1}{2} a(\mu; v, v) - f(\mu; v), \tag{1}$$

where the admissible set is defined as

$$\mathcal{K}(\mu) := \{ v \in \mathcal{V} \ \mid b(\mu; v, \eta) \leq g(\mu; \eta), \quad \forall \eta \in \mathcal{W}^+ \}, \tag{2}$$

and is assumed to be non-empty for all $\mu \in \mathcal{D}$.

We denote the boundedness and inf-sup stability coefficients of the bilinear form $b(\mu; \cdot, \cdot)$ with respect to the high-fidelity pair $(\mathcal{V}, \mathcal{W}^+)$ as follows:

$$c_{\mathrm{HF}}(\mu) := \sup_{\eta \in \mathcal{W}^+} \sup_{v \in \mathcal{V}} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}}, \qquad \beta_{\mathrm{HF}}(\mu) := \inf_{\eta \in \mathcal{W}^+} \sup_{v \in \mathcal{V}} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}}. \tag{3}$$

By the above assumption on $b(\mu; \cdot, \cdot)$, there exists a real number C_0 such that $c_{\mathrm{HF}}(\mu) \leq C_0$ for all $\mu \in \mathcal{D}$, and we additionally assume that the HF pair $(\mathcal{V}, \mathcal{W}^+)$ is such that

$$\exists \beta_0 > 0, \quad \forall \mu \in \mathcal{D}, \quad \beta_{\mathrm{HF}}(\mu) \geq \beta_0 > 0. \tag{4}$$

One possibility to solve the optimization problem (1) is to use a dual formulation. Let $\mathcal{L}(\mu; \cdot, \cdot) : \mathcal{V} \times \mathcal{W}^+ \to \mathbb{R}$ be the Lagrangian associated with (1) and defined as follows: For all $\mu \in \mathcal{D}$,

$$\mathcal{L}(\mu; v, \eta) := \frac{1}{2} a(\mu; v, v) - f(\mu; v) + b(\mu; v, \eta) - g(\mu; \eta), \quad \forall \ (v, \eta) \in \mathcal{V} \times \mathcal{W}^+. \tag{5}$$

Then, for all $\mu \in \mathcal{D}$, we can rewrite (1) as the following saddle-point problem: Find $(u(\mu), \lambda(\mu)) \in \mathcal{V} \times \mathcal{W}^+$ such that

$$(u(\mu), \lambda(\mu)) = \ \arg\min_{v \in \mathcal{V}} \max_{\eta \in \mathcal{W}^+} \mathcal{L}(\mu; v, \eta). \tag{6}$$

Owing to the above assumptions (in particular (4) and the coercivity of $a(\mu; \cdot, \cdot)$), the pair $(u(\mu), \lambda(\mu)) \in \mathcal{V} \times \mathcal{W}^+$ is uniquely defined and can be found as the critical point of the Lagrangian $\mathcal{L}(\mu; \cdot, \cdot)$ by solving for all $\mu \in \mathcal{D}$,

$$\begin{cases} a(\mu; u(\mu), v) + b(\mu; v, \lambda(\mu)) = f(\mu; v), & \forall v \in \mathcal{V}, \\ \qquad\qquad\qquad\qquad b(\mu; u(\mu), \eta) \leq g(\mu; \eta), & \forall \eta \in \mathcal{W}^+. \end{cases} \tag{7}$$

2.2 Decorrelated model reduction

Let $P \in \mathbb{N}^*$ and consider a family $\{(u(\mu_p), \lambda(\mu_p))\}_{p\in\{1:P\}} \subset \mathcal{V} \times \mathcal{W}^+$ of solutions to (7) which are computed by using a training subset $\mathcal{D}_{\texttt{train}} := \{\mu_p\}_{p\in\{1:P\}} \subset \mathcal{D}$ of cardinality P. In practice, the sampling of the parameter domain can be driven by a posteriori error estimates, as in [22, 5, 40]. Let us set

$$V := \boldsymbol{Span}(\{u(\mu_p)\}_{p\in\{1:P\}}) \subset \mathcal{V}, \qquad W^+ := \boldsymbol{Span}^+(\{\lambda(\mu_p)\}_{p\in\{1:P\}}) \subset \mathcal{W}^+, \tag{8}$$

where, for an arbitrary family $\{\theta_q\}_{q\in\{1:Q\}} \subset \mathcal{W}^+$ with $Q \in \mathbb{N}^*$, $\boldsymbol{Span}^+$ denotes the positive cone generated by setting

$$\boldsymbol{Span}^+(\{\theta_q\}_{q\in\{1:Q\}}) := \Big\{ \sum_{q\in\{1:Q\}} \alpha_q \theta_q,\ (\alpha_1, \cdots, \alpha_Q) \in \mathbb{R}^Q_+ \Big\}. \tag{9}$$

Given a positive real number $\delta_{\texttt{POD}} > 0$, one can construct using POD an orthonormal family of $N \in \mathbb{N}^*$ $(N \leq P)$ elements of V,

$$\{v_n\}_{n\in\{1:N\}} := \texttt{POD}(\{u(\mu_p)\}_{p\in\{1:P\}};\ \mathcal{V},\ \delta_{\texttt{POD}}),$$

such that

$$V_N := \boldsymbol{Span}(\{v_n\}_{n\in\{1:N\}}) \subset V, \tag{10}$$

$$e_{\texttt{POD}}(N) := \frac{\Big(\sum_{p\in\{1:P\}} \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N})(u(\mu_p))\|^2_{\mathcal{V}} \Big)^{\frac{1}{2}}}{\Big(\sum_{p\in\{1:P\}} \|u(\mu_p)\|^2_{\mathcal{V}} \Big)^{\frac{1}{2}}} \leq \delta_{\texttt{POD}}, \tag{11}$$

where $\Pi^{\mathcal{H}}_Z$ denotes the projection onto a generic closed convex subset Z of the generic Hilbert space $\mathcal{H}$ ($\Pi^{\mathcal{H}}_Z$ is the orthogonal projection if $\mathcal{Z}$ is a linear subspace) and $\mathbb{I}^{\mathcal{H}}$ denotes the identity operator in $\mathcal{H}$. Moreover, given a positive real number $\delta_{\texttt{CPG}} > 0$, one can construct using the CPG algorithm a subset $\{\chi_r\}_{r\in\{1:R\}}$ of $\{\lambda(\mu_p)\}_{p\in\{1:P\}}$ composed of $R \in \mathbb{N}^*$ $(R \leq P)$ vectors of W^+,

$$\{\chi_r\}_{r\in\{1:R\}} := \texttt{CPG}(\{\lambda(\mu_p)\}_{p\in\{1:P\}};\ \mathcal{W},\ \delta_{\texttt{CPG}}),$$

satisfying

$$W^+_R := \boldsymbol{Span}^+(\{\chi_r\}_{r\in\{1:R\}}) \subset W^+, \tag{12}$$

$$e_{\texttt{CPG}}(R) := \frac{\max_{p\in\{1:P\}} \|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{W^+_R})(\lambda(\mu_p))\|_{\mathcal{W}}}{\max_{p\in\{1:P\}} \|\lambda(\mu_p)\|_{\mathcal{W}}} \leq \delta_{\texttt{CPG}}. \tag{13}$$

For all $\mu \in \mathcal{D}$ and $N, R \leq P$, we define

$$\beta^{\text{dec}}_{N,R}(\mu) := \inf_{\eta\in W^+_R} \sup_{v\in V_N} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}}, \tag{14}$$

where the superscript refers to the decorrelated construction of the reduced bases. Since the primal reduced space V_N and the dual reduced cone W^+_R are constructed in a decorrelated manner, one cannot guarantee that $\beta^{\text{dec}}_{N,R}(\mu) > 0$ for all $\mu \in \mathcal{D}$, *i.e.*, the pair (V_N, W^+_R) may not satisfy an inf-sup condition.

2.3 Online enhancement of the reduced primal basis

A strategy for enriching the reduced primal basis in order to achieve inf-sup stability has been proposed in [22]. The idea is to complete the reduced primal basis $\{v_n\}_{n\in\{1:N\}}$ with as many functions as there are in the reduced dual basis $\{\chi_r\}_{r\in\{1:R\}}$. To this purpose, one proceeds as follows: For all μ in $\mathcal{D}$, one defines the operator $\mathbb{T}(\mu) : \mathcal{W}^+ \to \mathcal{V}$ such that

$$\mathbb{T}(\mu) := J \circ \mathcal{B}_{\mathrm{HF}}(\mu), \tag{15}$$

where $J : \mathcal{V}' \to \mathcal{V}$ is the Riesz isomorphism between $\mathcal{V}'$ and $\mathcal{V}$ and where the operator $\mathcal{B}_{\mathrm{HF}}(\mu) : \mathcal{W}^+ \to \mathcal{V}'$ is defined such that

$$\langle \mathcal{B}_{\mathrm{HF}}(\mu)(\eta), v\rangle_{\mathcal{V}',\mathcal{V}} := b(\mu; v, \eta), \quad \forall (v,\eta) \in \mathcal{V} \times \mathcal{W}^+. \tag{16}$$

Then, for all $\mu \in \mathcal{D}$, the enriched reduced primal space $V^{\mathrm{on}}_{N,R}(\mu)$ is defined as

$$V^{\mathrm{on}}_{N,R}(\mu) := V_N + S_R(\mu) \subset \mathcal{V}, \quad S_R(\mu) := \boldsymbol{Span}(\{\mathbb{T}(\mu)\chi_r\}_{r\in\{1:R\}}). \tag{17}$$

The superscript refers to the online construction of the enriched basis. The main motivation for (17) is that it ensures that the bilinear form $b(\mu; \cdot, \cdot)$ is uniformly inf-sup stable with respect to the pair $(V^{\mathrm{on}}_{N,R}(\mu), W^+_R)$, *i.e.*, we have

$$\forall \mu \in \mathcal{D}, \quad \beta^{\mathrm{on}}_{N,R}(\mu) := \inf_{\eta\in W^+_R} \sup_{v\in V^{\mathrm{on}}_{N,R}(\mu)} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}}\|\eta\|_{\mathcal{W}}} \geq \beta_{\mathrm{HF}}(\mu) \geq \beta_0 > 0. \tag{18}$$

Indeed, for all $\eta \in W^+_R$, we have

$$\begin{aligned}\sup_{v\in V^{\mathrm{on}}_{N,R}(\mu)} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}}} &\geq \frac{b(\mu; \mathbb{T}(\mu)(\eta), \eta)}{\|\mathbb{T}(\mu)(\eta)\|_{\mathcal{V}}} = \frac{\|\mathbb{T}(\mu)(\eta)\|^2_{\mathcal{V}}}{\|\mathbb{T}(\mu)(\eta)\|_{\mathcal{V}}} = \|\mathbb{T}(\mu)(\eta)\|_{\mathcal{V}} \\ &= \|J(\mathcal{B}_{\mathrm{HF}}(\mu)(\eta))\|_{\mathcal{V}} = \|\mathcal{B}_{\mathrm{HF}}(\mu)(\eta)\|_{\mathcal{V}'} \geq \beta_{\mathrm{HF}}(\mu)\|\eta\|_{\mathcal{W}}.\end{aligned} \tag{19}$$

In the case where the operator $\mathcal{B}_{\mathrm{HF}}$ is parameter-independent, the space $V^{\mathrm{on}}_{N,R}$ is also parameter-independent. This space can thus be constructed once and for all in the offline phase, as in [22]. However, when the operator $\mathcal{B}_{\mathrm{HF}}(\mu)$ is parameter-dependent, the enriched space $V^{\mathrm{on}}_{N,R}(\mu)$ shares the same feature, and thus has to be constructed in the online phase, which is computationally inefficient. We overcome this problem in the next section.

3 Computationally efficient, stable model reduction

In this section, the reduced subspaces V_N and W^+_R obtained respectively by (10) and (12) are fixed. Our goal is to construct a parameter-independent subspace of $\mathcal{V}$ that provides a sufficiently accurate approximation of $V^{\mathrm{on}}_{N,R}(\mu)$ for all $\mu \in \mathcal{D}$ to preserve inf-sup stability. The crucial advantage is that this parameter-independent subspace can be constructed once and for all in the offline phase. The idea is to approximate the linear space

$$S_R := \underset{\mu\in\mathcal{D}_{\mathtt{train}}}{+} S_R(\mu), \tag{20}$$

by a subspace $S^{\mathrm{red}}_R \subset S_R$, so that the bilinear form $b(\mu; \cdot, \cdot)$ is inf-sup stable with respect to the pair $(V_N + S^{\mathrm{red}}_R, W^+_R)$ for all $\mu \in \mathcal{D}_{\mathtt{train}}$. We will see in our numerical experiments that it is possible to achieve this property by a proper subspace $S^{\mathrm{red}}_R \subset S_R$ having a much smaller dimension than S_R. To realize this strategy, we rely on the following theoretical result which gives a sufficient condition on a subspace S of S_R to guarantee the above inf-sup stability.

Proposition 3.1 (Inf-sup stability)**.** *Let S be any finite-dimensional subspace of S_R. For all $\mu \in \mathcal{D}$, we define*

$$\sigma_S(\mu) := \sup_{\substack{v \in S_R(\mu) \\ \|v\|_{\mathcal{V}} \leq 1}} \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N+S})(v)\|_{\mathcal{V}} = \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N+S})|_{S_R(\mu)}\|_{\mathcal{L}(\mathcal{V})}, \tag{21}$$

and the boundedness constant

$$c_S(\mu) := \sup_{\eta \in W_R^+} \sup_{v \in V_N + S_R(\mu) + S} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}} \leq C_0. \tag{22}$$

For all $\mu \in \mathcal{D}$, if

$$\sigma_S(\mu) < \frac{\beta^{\mathrm{on}}_{N,R}(\mu)}{c_S(\mu)}, \tag{23}$$

then the following inf-sup condition holds:

$$\inf_{\eta \in W_R^+} \sup_{v \in V_N + S} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}} \geq \beta^*_S(\mu) := \frac{\beta^{\mathrm{on}}_{N,R}(\mu) - c_S(\mu)\sigma_S(\mu)}{1 + \sigma_S(\mu)} > 0. \tag{24}$$

Proof. Let $\mu \in \mathcal{D}$ and let $\eta \in W_R^+$ with $\|\eta\|_{\mathcal{W}} = 1$. The inf-sup stability of $b(\mu; \cdot, \cdot)$ with respect to the pair $(V^{\mathrm{on}}_{N,R}(\mu), W_R^+)$ (see (18)) implies that there exists $v_\eta \in S_R(\mu)$ such that

$$b(\mu; v_\eta, \eta) \geq 1, \qquad \beta^{\mathrm{on}}_{N,R}(\mu)\|v_\eta\|_{\mathcal{V}} \leq 1. \tag{25}$$

Letting $u_\eta := \Pi^{\mathcal{V}}_{V_N+S}(v_\eta)$, the definition of $\sigma_S(\mu)$ implies that

$$\|v_\eta - u_\eta\|_{\mathcal{V}} = \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N+S})(v_\eta)\|_{\mathcal{V}} \leq \sup_{\substack{v \in S_R(\mu) \\ \beta^{\mathrm{on}}_{N,R}(\mu)\|v\|_{\mathcal{V}} \leq 1}} \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N+S})(v)\|_{\mathcal{V}} = \frac{\sigma_S(\mu)}{\beta^{\mathrm{on}}_{N,R}(\mu)}. \tag{26}$$

On the one hand, we have

$$b(\mu; u_\eta, \eta) = b(\mu; u_\eta - v_\eta, \eta) + b(\mu; v_\eta, \eta),$$

and the definition (22) of $c_S(\mu)$ implies that

$$|b(\mu; u_\eta - v_\eta, \eta)| \leq c_S(\mu)\|u_\eta - v_\eta\|_{\mathcal{V}}\|\eta\|_{\mathcal{W}} = c_S(\mu)\|u_\eta - v_\eta\|_{\mathcal{V}}.$$

Hence, recalling that $b(\mu; v_\eta, \eta) \geq 1$ owing to (25), we have

$$\begin{aligned} b(\mu; u_\eta, \eta) &= b(\mu; u_\eta - v_\eta, \eta) + b(\mu; v_\eta, \eta) \\ &\geq 1 - c_S(\mu)\|u_\eta - v_\eta\|_{\mathcal{V}} \geq 1 - \frac{c_S(\mu)}{\beta^{\mathrm{on}}_{N,R}(\mu)}\sigma_S(\mu), \end{aligned}$$

where the last bound follows from the inequality (26). On the other hand, we have

$$\|u_\eta\|_{\mathcal{V}} \leq \|v_\eta - u_\eta\|_{\mathcal{V}} + \|v_\eta\|_{\mathcal{V}} \leq \frac{\sigma_S(\mu)}{\beta^{\mathrm{on}}_{N,R}(\mu)} + \frac{1}{\beta^{\mathrm{on}}_{N,R}(\mu)} = \frac{1 + \sigma_S(\mu)}{\beta^{\mathrm{on}}_{N,R}(\mu)}.$$

We conclude that

$$\frac{b(\mu; u_\eta, \eta)}{\|u_\eta\|_{\mathcal{V}}} \geq \frac{\beta^{\mathrm{on}}_{N,R}(\mu)}{1 + \sigma_S(\mu)}\left(1 - \frac{c_S(\mu)}{\beta^{\mathrm{on}}_{N,R}(\mu)}\sigma_S(\mu)\right) = \frac{\beta^{\mathrm{on}}_{N,R}(\mu) - c_S(\mu)\sigma_S(\mu)}{1 + \sigma_S(\mu)} > 0,$$

where the last bound follows from the assumption (23). Since $u_\eta \in V_N + S$, this implies the inf-sup condition (24). □

Using the result of Prop. 3.1, we design in Algo. 1 a so-called Projected Greedy Algorithm (PGA) that, given a training subset $\mathcal{D}_{\texttt{train}} \subset \mathcal{D}$, the vector space V_N, and a tolerance $\delta > 0$, returns a subspace $S_R^{\text{red}} \subset S_R$ such that

$$\sup_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S_R^{\text{red}}}(\mu) \leq \delta. \tag{27}$$

Notice that since $\sigma_{S_R}(\mu) = 0$ for all $\mu \in \mathcal{D}$, it is possible to construct a subspace S_R^{red} of S_R satisfying (27). Our numerical results presented in Section 5 indicate that it is reasonable to expect that S_R^{red} has a (much) smaller dimension than S_R. In practice, the space S_R^{red} is constructed in a progressive way by means of a greedy algorithm. The PGA algorithm involves the following two main steps at iteration n:

- seek the parameter μ_n by solving an eigenvalue problem (line 2 and line 9).
- seek the supremizer $v_{n+1}^{(1)} \in S_R(\mu_n)$ by solving a linear system (line 5).

Algorithm 1 `PGA: Projected Greedy Algorithm` | $\texttt{PGA}(\mathcal{D}_{\texttt{train}}, V_N, \delta)$

Require: $\mathcal{D}_{\texttt{train}} \subset \mathcal{D}$: `training subset`
$V_N \subset \mathcal{V}$: `primal reduced space`
$\delta > 0$: `tolerance`
Ensure: S: `subspace of` S_R

1: $S^0 := \{0\}$
2: $\mu_0 \in \underset{\mu \in \mathcal{D}_{\texttt{train}}}{\texttt{argmax}}\ \sigma_{S^0}(\mu)$
3: $n := 0$
4: **while** $\sigma_{S^n}(\mu_n) > \delta$ **do**
5: $\quad v_{n+1}^{(1)} \in \underset{\substack{v \in S_R(\mu_n) \\ \|v\|_{\mathcal{V}} \leq 1}}{\texttt{argmax}}\ \|(\mathbb{I}^{\mathcal{V}} - \Pi_{V_N + S^n}^{\mathcal{V}})(v)\|_{\mathcal{V}}$
6: $\quad v_{n+1}^{(2)} := (\mathbb{I}^{\mathcal{V}} - \Pi_{V_N + S^n}^{\mathcal{V}})(v_{n+1}^{(1)})$
7: $\quad v_{n+1} := \frac{v_{n+1}^{(2)}}{\|v_{n+1}^{(2)}\|_{\mathcal{V}}}$
8: $\quad S^{n+1} := S^n + \boldsymbol{Span}\{v_{n+1}\}$
9: $\quad \mu_{n+1} \in \underset{\mu \in \mathcal{D}_{\texttt{train}}}{\texttt{argmax}}\ \sigma_{S^{n+1}}(\mu)$
10: $\quad n = n + 1$
11: **end while**
12: **return** $S := S^n$

In conclusion, we define the enriched primal reduced space $V_{N,R}^{\text{off}}$ as follows:

$$S_R^{\text{red}} := \texttt{PGA}(\mathcal{D}_{\texttt{train}}, V_N, \delta_{\texttt{PGA}}) \subset S_R \subset \mathcal{V}, \tag{28}$$

$$V_{N,R}^{\text{off}} := V_N + S_R^{\text{red}} \subset \mathcal{V}, \tag{29}$$

where the tolerance $\delta_{\texttt{PGA}} > 0$ is small enough so that (23) is satisfied for all $\mu \in \mathcal{D}_{\texttt{train}}$. In practice, a simple choice is $\delta_{\texttt{PGA}} < \beta_0^{-1} C_0$ since $\beta_{N,R}^{\text{on}}(\mu) \geq \beta_0$ and $c_S(\mu) \leq C_0$ for all $\mu \in \mathcal{D}_{\texttt{train}}$. The superscript in the notation $V_{N,R}^{\text{off}}$ refers to the offline construction of the enriched reduced primal basis. For all

$\mu \in \mathcal{D}$, we define

$$\beta^{\text{off}}_{N,R}(\mu) := \inf_{\eta \in W_R^+} \sup_{v \in V^{\text{off}}_{N,R}} \frac{b(\mu; v, \eta)}{\|v\|_{\mathcal{V}} \|\eta\|_{\mathcal{W}}}. \tag{30}$$

Notice that we only guarantee that $\beta^{\text{off}}_{N,R}(\mu) > 0$ for all $\mu \in \mathcal{D}_{\texttt{train}}$. It is reasonable to expect inf-sup stability for all $\mu \in \mathcal{D}$ if the training subset is sufficiently rich.

Let us now show that Algo. 1 necessarily terminates in a finite number of iterations.

Lemma 3.2 (Convergence of PGA algorithm)**.** *The sequence* $\left(\max_{\mu \in \mathcal{D}_{train}} \sigma_{S^n}(\mu)\right)_{n \geq 0}$ *is (non-strictly) decreasing. Moreover, there is* $n_0 \leq \min(m_{\mathcal{D}_{train}} R, \dim(\mathcal{V}) - \dim(V_N))$ *with* $m_{\mathcal{D}_{train}} := \#(\mathcal{D}_{train})$ *such that*

$$\max_{\mu \in \mathcal{D}_{train}} \sigma_{S^n}(\mu) = 0, \qquad \forall n \geq n_0. \tag{31}$$

Proof. (1) Line 8 of Algo. 1 implies that $S^n \subseteq S^{n+1}$ for all $n \geq 0$. Hence, for all $\mu \in \mathcal{D}_{\texttt{train}}$,

$$\max_{v \in S_R(\mu)} \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N + S^{n+1}})(v)\|_{\mathcal{V}} \leq \max_{v \in S_R(\mu)} \|(\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N + S^{n}})(v)\|_{\mathcal{V}},$$

which implies that

$$\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^{n+1}}(\mu) \leq \max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu).$$

This proves the first assertion.

(2) Let us now prove the second assertion. To this purpose, let us first prove by induction that for all $n \geq 0$, if $\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu) > 0$, then $S^n \subsetneq S^{n+1}$ and

$$\dim\,(V_N + S^{n+1}) = \dim\,(V_N + S^n) + 1. \tag{32}$$

Indeed, since $\mu_n \in \texttt{argmax}_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu)$, we have $\sigma_{S^n}(\mu_n) = \max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu) > 0$. Hence, $v^{(1)}_{n+1}$ from line 5 satisfies

$$v^{(2)}_{n+1} := (\mathbb{I}^{\mathcal{V}} - \Pi^{\mathcal{V}}_{V_N + S^n})(v^{(1)}_{n+1}) \neq 0.$$

Thus, $v^{(2)}_{n+1}$ is nonzero and orthogonal to the space $V_N + S^n$. Since $S^{n+1} := S^n + \boldsymbol{Span}\{v_{n+1}\}$, this proves (32). We are now in a position to prove the second assertion. We first observe that if there is $n_0 \geq 0$ such that $\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^{n_0}}(\mu) = 0$, then $\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu) = 0$ for all $n \geq n_0$. Moreover, since $S^n \subseteq S_R$ and $\dim(S_R) \leq m_{\mathcal{D}_{\texttt{train}}} R$, we infer from (21) that $\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu) = 0$ whenever $n \geq m_{\mathcal{D}_{\texttt{train}}} R$. Finally, reasoning by induction and using (32) shows that $\dim\,(V_N + S^n) = \dim(V_N) + n$ for all $n \geq 0$ such that $\max_{\mu \in \mathcal{D}_{\texttt{train}}} \sigma_{S^n}(\mu) > 0$. Since $V_N + S^n \subset \mathcal{V}$, we must have $n \leq \dim(\mathcal{V}) - \dim(V_N)$. This completes the proof of (31). □

Remark 3.3 (Finite termination)*.* Lemma 3.2 provides two upper bounds on the maximum number of iterations of the PGA algorithm. The first upper bound, $n_0 \leq m_{\mathcal{D}_{\texttt{train}}} R$ is, in particular, independent of the dimension of the HF space $\mathcal{V}$.

Remark 3.4 (Comparison with [14])*.* The PGA algorithm is similar in spirit to the Algorithms 2 and 3 proposed in [14] in the context of Petrov–Galerkin reduced basis approximations of transport equations (the adaptation to saddle-point problems is discussed in Section 7 therein). The overall organization of the reduced basis method is, however, different. Indeed, in contrast to [14] where a double greedy algorithm is proposed, we do not enrich here the reduced dual space in a greedy way and adapt the reduced primal space to guarantee inf-sup stability at each iteration of the greedy algorithm. Instead, we fix once and for all the reduced dual space (which can be obtained as the output of a greedy algorithm, see Section 4), and then enrich the reduced primal space following the strategy highlighted in Algo. 1. A computational comparison of the performances of both algorithms is postponed to future work. Moreover, we notice that Proposition 3.1 relates the inf-sup constant of the pair of reduced spaces to some approximation property of the subspace $V_N + S$ in $\mathcal{V}$. This idea is similar to [14, Prop. 3.7] which relates inf-sup stability to the notion of δ-proximality.

4 Modified CPG algorithm

In this section, we propose a modified version of the plain CPG algorithm (see Remark 4.4 for a comparison with the algorithm from [6]). The goal is to promote the generation of a cone having a larger aperture than the one built by the plain CPG algorithm. Indeed, a problem that arises numerically when working with a positive cone is to be able to generate a basis such that the corresponding Gram matrix is as well-conditioned as possible. In practice, one has often to deal with bases composed of vectors which are almost collinear. In this context, ill-conditioning issues can arise, in particular where performing projections onto the cone. To overcome this difficulty, it is not possible to perform an orthogonalization process (such as Gram–Schmidt) since this would lead to a departure from the positive cone $\mathcal{W}^+$. Instead, one possibility is to promote the aperture of the cone produced by the plain CPG algorithm.

Specifically, we introduce the mCPG (modified CPG) algorithm (see Algo. 2) which, given a set of $Q \in \mathbb{N}^*$ vectors $\{\theta_q\}_{q\in\{1:Q\}}$ of $\mathcal{W}^+$, and a positive real number δ, produces a set of $R \in \mathbb{N}^*(R \leq Q)$ vectors of $\mathcal{W}^+$,

$$\{\nu_r\}_{r\in\{1:R\}} := \texttt{mCPG}(\{\theta_q\}_{q\in\{1:Q\}};\ \mathcal{W},\ \delta),$$

satisfying

$$W_R^+ := \boldsymbol{Span}^+(\{\nu_r\}_{r\in\{1:R\}}) \subset \mathcal{W}^+, \tag{33}$$

$$e_{\texttt{mCPG}}(R) := \frac{\max\limits_{q\in\{1:Q\}} \|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{W_R^+})(\theta_q)\|_{\mathcal{W}}}{\max\limits_{q\in\{1:Q\}} \|\theta_q\|_{\mathcal{W}}} \leq \delta. \tag{34}$$

Remark 4.1 (Cone aperture). For all $n \in \mathbb{N}^*$, we have $\boldsymbol{Span}^+(\{\theta_{q_n}\}_{n\in\{1:r\}}) \subset \boldsymbol{Span}^+(\{\nu_n\}_{n\in\{1:r\}})$. Indeed, reasoning by induction, since we have $\nu_r := \frac{\theta_{q_r} - \Upsilon_r}{\|\theta_{q_r} - \Upsilon_r\|_{\mathcal{W}}}$ with $\Upsilon_r \in K_{r-1}$, we infer that $\theta_{q_r} = \|\theta_{q_r} - \Upsilon_r\|_{\mathcal{W}} \times \nu_r + \Upsilon_r \in \boldsymbol{Span}^+(\{\nu_n\}_{n\in\{1:r\}})$.

Remark 4.2 (Proper termination). In the mCPG algorithm, we observe that line 10 is only executed if $\nu_r \neq 0$. Indeed, if $\nu_r = 0$, then $\theta_{q_r} = \Upsilon_r \in K_{r-1}$ and consequently $e_{r-1} = \frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{K_{r-1}})(\theta_{q_r})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}} = 0 < \delta$. This means that we do not continue the loop since the condition $e_{r-1} > \delta$ is not satisfied.

Remark 4.3 (Comparison with CPG). The main difference between the CPG and mCPG algorithms is that at each iteration $r \geq 1$, the CPG algorithm does not execute line 9 and simply sets $\nu_r = \frac{\theta_{q_r}}{\|\theta_{q_r}\|_{\mathcal{W}}}$. Instead, the mCPG algorithm computes ν_r as a member of $\mathcal{W}^+$ (in fact of $\boldsymbol{Span}^+(\{\theta_{q_n}\}_{n\in\{1:r\}})$). We will show in the numerical experiments of Section 5 that the mCPG algorithm has the advantage of providing a set of vectors whose Gram matrix is better conditioned compared to the one obtained with the CPG algorithm. This property comes from the fact that at iteration r of the mCPG algorithm, the element r of the basis is constructed by removing all the information contained in the basis constructed at iteration $r-1$. Figure 1 presents an illustration of how the CPG and mCPG algorithms operate on the first two iterations.

Remark 4.4 (Comparison with [6]). The algorithm proposed in [6] consists in constructing from a given family of functions $\{\nu_r\}_{r\in\{1:R\}}$ (which can be obtained, for instance, as the output of the CPG algorithm from an original family of functions $\{\theta_q\}_{q\in\{1:Q\}}$), a new family of modified functions $\{\widetilde{\nu}_r\}_{r\in\{1:R\}}$ so that

$$\boldsymbol{Span}^+(\{\nu_r\}_{r\in\{1:R\}}) \subset \boldsymbol{Span}^+(\{\widetilde{\nu}_r\}_{r\in\{1:R\}}).$$

In other words, the approach proposed in [6] allows one to construct an enlarged cone from a reduced cone which has to be given as an input, and which is spanned by the same number of functions as

Algorithm 2 `mCPG: modified CPG algorithm` | $\texttt{mCPG}(\{\theta_q\}_{q\in\{1:Q\}}; \mathcal{W}, \delta)$

Require: $\{\theta_q\}_{q\in\{1:Q\}}, Q \in \mathbb{N}^*$: `set of vectors`
$\mathcal{W}$: `Hilbert space`
$\delta > 0$: `tolerance`
Ensure: $\{\nu_r\}_{r\in\{1:R\}} \subset \mathcal{W}^+, R \leq Q$

$r := 0$
$I_0 := \emptyset$
$K_0 := \{0\} \subset \mathcal{W}^+$
$e_0 := 1 + \delta$
$q_1 := \underset{q\in\{1:Q\}}{\texttt{argmax}} \|\theta_q\|_{\mathcal{W}}$
while $(e_r > \delta)$ **and** $(r < Q)$ **do**
  $r := r + 1$
  $I_r := I_{r-1} \cup \{q_r\}$
  $\Upsilon_r := \underset{\Upsilon\in K_{r-1}\cap(\theta_{q_r}-\mathcal{W}^+)}{\texttt{argmin}} \|\theta_{q_r} - \Upsilon\|_{\mathcal{W}}$
  $\nu_r := \frac{\theta_{q_r}-\Upsilon_r}{\|\theta_{q_r}-\Upsilon_r\|_{\mathcal{W}}}$
  $K_r := K_{r-1} + \boldsymbol{Span}^+\{\nu_r\}$
  $q_{r+1} := \underset{q\in\{1:Q\}\setminus I_r}{\texttt{argmax}} \frac{\|(\mathbb{I}^{\mathcal{W}}-\Pi^{\mathcal{W}}_{K_r})(\theta_q)\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}}$
  $e_r := \frac{\|(\mathbb{I}^{\mathcal{W}}-\Pi^{\mathcal{W}}_{K_r})(\theta_{q_{r+1}})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}}$
end while
$R := r$
return $\{\nu_n\}_{n\in\{1:R\}}$

the orginal reduced cone. In contrast, the mCPG algorithm proposed herein consists in constructing an enlarged cone directly from the original family of functions $\{\theta_q\}_{q\in\{1:Q\}}$ as an output of a greedy algorithm, by enlarging the current reduced cone at each step of the greedy algorithm. This has, in our opinion, two advantages. First, it avoids ill-conditioning issues raised by the standard CPG algorithm and which are then also encountered in the approach proposed in [6]. Second, considering enlarge reduced cones at each iteration of a greedy algorithm yields a tighter control on the accuracy of the results given by the corresponding reduced-order model. Finally, we observe that our numerical experiments indicate that the accuracy of the reduced-order approximation is better when using the mCPG algorithm than the CPG algorithm. A more detailed computational comparison with the approach from [6] is left to future work.

Finally, the following result states that Algo. 2 terminates in a finite number of iterations.

Lemma 4.5 (Finite termination)**.** *The sequence $(e_r)_{1\leq r\leq Q}$ is (non-strictly) decreasing and $e_Q = 0$.*

Proof. Let $1 \leq r \leq Q-1$ and let us prove that $e_{r+1} \leq e_r$. First, we have $K_r \subset K_{r+1}$. Hence,

$$e_{r+1} := \frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{K_{r+1}})(\theta_{q_{r+2}})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}} \leq \frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{K_r})(\theta_{q_{r+2}})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}}.$$

In addition, since

$$q_{r+1} := \underset{q\in\{1:Q\}\setminus I_r}{\texttt{argmax}} \frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi^{\mathcal{W}}_{K_r})(\theta_q)\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}},$$

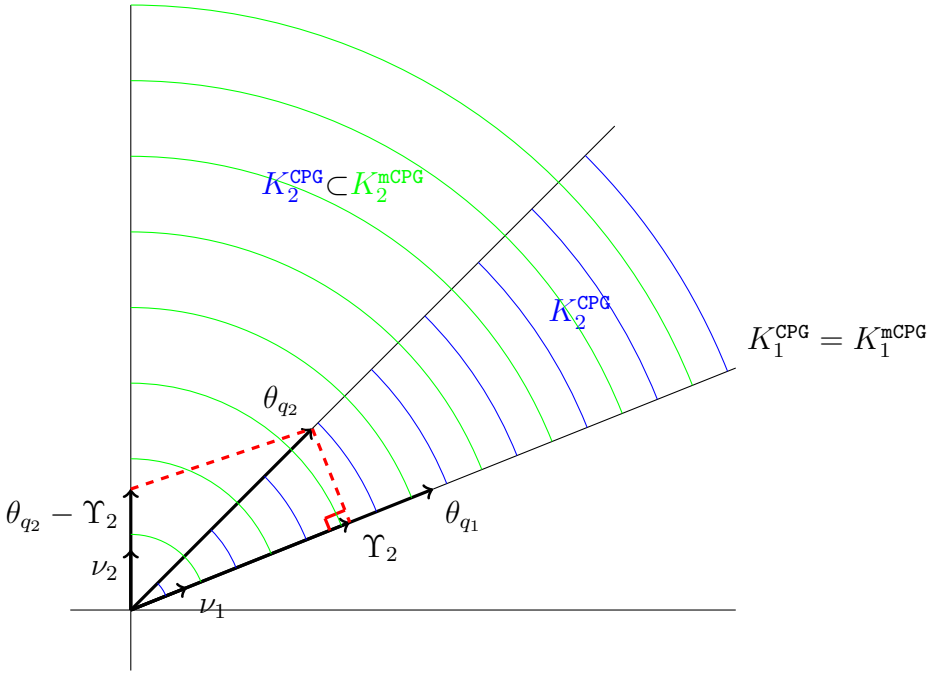


Figure 1: Comparison of the first two iterations of the CPG and mCPG algorithms.

we infer that

$$\frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi_{K_r}^{\mathcal{W}})(\theta_{q_{r+2}})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}} \leq \frac{\|(\mathbb{I}^{\mathcal{W}} - \Pi_{K_r}^{\mathcal{W}})(\theta_{q_{r+1}})\|_{\mathcal{W}}}{\|\theta_{q_1}\|_{\mathcal{W}}} =: e_r.$$

This proves that the sequence $(e_r)_{1\leq r\leq Q}$ is (non-strictly) decreasing. Finally, since necessarily $\boldsymbol{Span}^+(\{\theta_q\}_{q\in\{1:Q\}})$ K_Q, we infer that $e_Q = 0$. This completes the proof. □

5 Numerical results

In this section, we numerically illustrate our theoretical results on the membrane obstacle problem and on the frictionless Hertz contact problem. In both examples, we compare the computational performance of the different methods for building the reduced primal space (decorrelated, online and offline) and the reduced dual basis (CPG and mCPG). The HF computations use a combination of `Freefem++` [23] and `Python`, whereas the algorithms presented herein have been developed in `Python` using the convex optimization package `cvxopt` [4].

5.1 Membrane obstacle problem

In this section, we study the membrane obstacle problem. We consider a square, elastic membrane $\Omega \subset \mathbb{R}^2$ of side $A = 1$m, located at some distance from an obstacle represented by a circular sub-domain $\omega(\mu) \subset \Omega$ of radius $r := \mu_1$ and centre $c := (\mu_2, \mu_3)$ with

$$\mu = (\mu_1, \mu_2, \mu_3) \in \mathcal{D} := [0.8, 1.2] \times [-0.05, 0.05] \times [-0.05, 0.05](\mathrm{m}). \tag{35}$$

For all $\mu \in \mathcal{D}$, we assume that there is a smooth invertible geometric mapping $h(\mu) : \widehat{\Omega} := \Omega \longrightarrow \Omega$ such that there is a reference domain $\widehat{\omega} \subset \Omega$ satisfying $h(\mu)(\widehat{\omega}) = \omega(\mu)$. We set $h_1(\mu) := h(\mu)|_{\widehat{\omega}}$. Moreover, the function $\widehat{\psi} : \widehat{\omega} \to \mathbb{R}$ prescribes the elevation of the obstacle in the reference domain, and we set $\psi(\mu) := \widehat{\psi} \circ h_1(\mu)^{-1}$. We apply to the membrane a vertical load $\widehat{\ell} : \widehat{\Omega} \to \mathbb{R}$, with $\widehat{\ell} \in L^2(\widehat{\Omega})$

and we set $\ell(\mu) := \widehat{\ell} \circ h(\mu)^{-1}$. The membrane is fixed on the boundary Γ of Ω (see Figure 2). Thus, we consider the following model problem:

$$\begin{cases} -\Delta u = \ell(\mu), & \text{in } \Omega, \\ \quad u \geq \psi(\mu), & \text{in } \omega(\mu), \\ \quad u = 0, & \text{on } \Gamma. \end{cases} \tag{36}$$

In what follows, we take $\widehat{\psi}$ so that $\psi(\mu)(x,y) := -1.25\frac{(x-\mu_2)^2+(y-\mu_3)^2}{\mu_1^2}$ for all $\mu \in \mathcal{D}$ and all $(x,y) \in \omega(\mu)$, and $\widehat{\ell}(x,y) = -1$ for all $(x,y) \in \widehat{\Omega}$. The coefficient -1.25 in front of the obstacle function is chosen so that the constraints at the boundary of $\omega(\mu)$ are inactive in order to avoid oscillations of the Lagrange multiplier at the transition zone. The training set is chosen as $\mathcal{D}_{\text{train}} := \{0.8 + 0.1i, 0 \leq i \leq 4\} \times \{-0.05 + 0.025i, 0 \leq i \leq 4\}^2$ (altogether 125 points) and the validation set $\mathcal{D}_{\text{valid}}$ is generated by choosing 64 elements in $\mathcal{D}$ randomly with a uniform distribution.

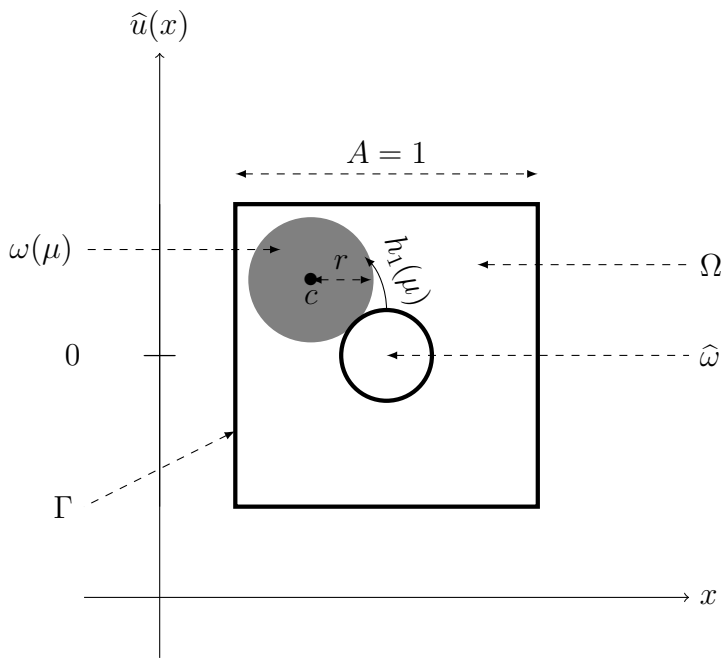


Figure 2: Obstacle test-case: Configuration.

We consider the HF subspace $\widehat{\mathcal{V}} \subset H_0^1(\widehat{\Omega}; \mathbb{R})$ and the HF subcone $\widehat{\mathcal{W}}^+ \subset L^2(\widehat{\omega}; \mathbb{R}_+)$ built using $\mathbb{P}_1$ finite elements for the displacement and the Lagrange multiplier [15, 10]. Notice that the displacement and the Lagrange multiplier are now defined on the reference domains $\widehat{\Omega}$ and $\widehat{\omega}$, respectively. The reference mesh of $\widehat{\Omega}$ is fitted to the boundary of $\widehat{\omega}$ and is composed of 3594 nodes with 467 nodes in $\widehat{\omega}$. Thus, we have $\mathcal{N} := \dim(\widehat{\mathcal{V}}) = 3594$ primal and $\mathcal{R} := \dim(\widehat{\mathcal{W}}) = 467$ dual degrees of freedom. For all $\mu \in \mathcal{D}$, the mesh of Ω fitted to the boundary of $\omega(\mu)$ is built using the geometric mapping $h(\mu)$. We have verified that all the meshes maintain satisfactory regularity properties by using the aspect ratio quality criterion from the Salome platform [35]. Figure 3 displays the maximum aspect ratio of the meshes for all $\mu \in \mathcal{D}_{\texttt{train}} \cup \mathcal{D}_{\texttt{valid}}$.

The variational formulation of (36) leads to a minimization problem of the form (1) with the bilinear form $\widehat{a}(\mu; \cdot, \cdot) : \widehat{\mathcal{V}} \times \widehat{\mathcal{V}} \to \mathbb{R}$ defined by

$$\widehat{a}(\mu; \widehat{u}, \widehat{v}) := \int_\Omega \nabla(\widehat{u} \circ h(\mu)^{-1}) \cdot \nabla(\widehat{v} \circ h(\mu)^{-1})\, d\Omega, \tag{37}$$

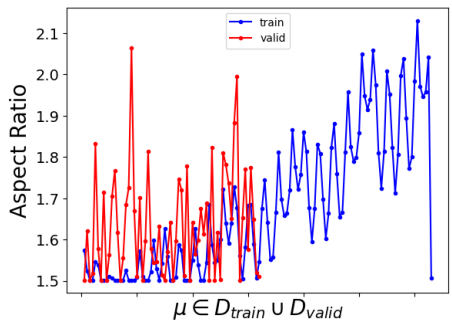


Figure 3: Obstacle test-case: Maximum aspect ratio for $\mu \in \mathcal{D}_{\texttt{train}} \cup \mathcal{D}_{\texttt{valid}}$.

and the linear form $\widehat{f}(\mu;\cdot) : \widehat{\mathcal{V}} \to \mathbb{R}$ defined by

$$\widehat{f}(\mu;\widehat{v}) := \int_{\Omega} (\widehat{\ell} \circ h(\mu)^{-1})(\widehat{v} \circ h(\mu)^{-1})\, d\Omega. \tag{38}$$

The admissible set $\widehat{\mathcal{K}}(\mu)$ associated with the obstacle condition is defined as follows:

$$\widehat{\mathcal{K}}(\mu) := \{\widehat{v} \in \widehat{\mathcal{V}} \ \mid \widehat{b}(\mu;\widehat{v},\widehat{\eta}) \leq \widehat{g}(\mu;\widehat{\eta}),\ \forall \widehat{\eta} \in \widehat{\mathcal{W}}^{+}\}, \tag{39}$$

where

$$\widehat{b}(\mu;\widehat{v},\widehat{\eta}) := \int_{\omega(\mu)} (\widehat{v} \circ h_1(\mu)^{-1})(\widehat{\eta} \circ h_1(\mu)^{-1})\, d\Gamma, \qquad \widehat{g}(\mu;\widehat{\eta}) := \int_{\omega(\mu)} (\widehat{\psi} \circ h_1(\mu)^{-1})(\widehat{\eta} \circ h_1(\mu)^{-1})\, d\Gamma. \tag{40}$$

Figure 4 displays the deformed configuration resulting from the HF displacement field $u(\mu) := \widehat{u}(\mu) \circ h(\mu)^{-1}$ (left panel) and the Lagrange multiplier $\lambda(\mu) := \widehat{\lambda}(\mu) \circ h(\mu)^{-1}$ (right panel) for $\mu = (0.18, 0.025, -0.025)$(m). Figure 5 shows on its left (resp. right) panel the projection error $\widehat{e}_{\texttt{POD}}$

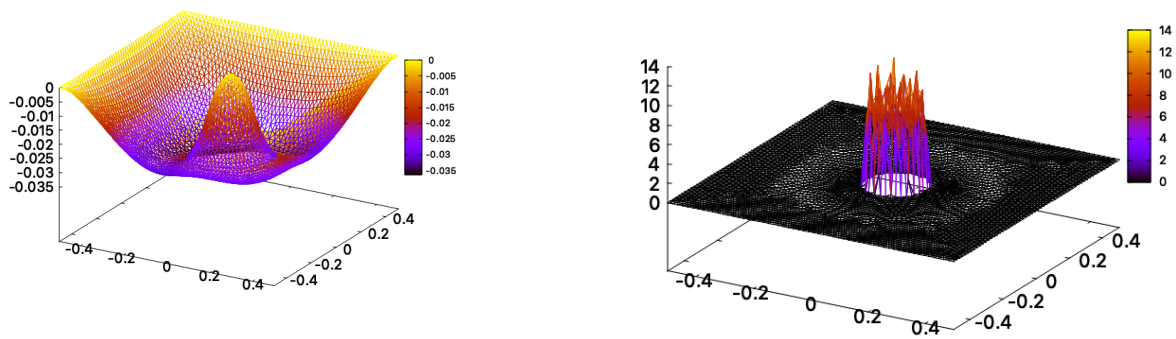


Figure 4: Obstacle test-case: HF solution for $\mu = (0.18, 0.025, -0.025)$. Left: $u(\mu)$. Right: $\lambda(\mu)$.

(resp. $\widehat{e}_{\texttt{CPG}}$ and $\widehat{e}_{\texttt{mCPG}}$) defined in (11) (resp. (13) and (34)) produced by the POD (resp. CPG and mCPG) algorithms as a function of the number of vectors composing the reduced primal (resp. dual) bases. In all cases, the projection errors decrease sufficiently fast so that the linear spaces generated by the primal and dual snapshots can be approximated by small-dimensional subspaces. We also notice

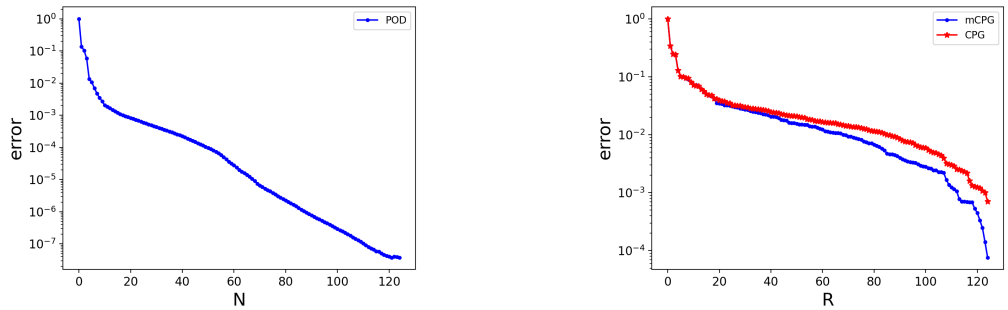


Figure 5: Obstacle test-case: projection errors for the POD, CPG and mCPG algorithms. Left: POD. Right: CPG and mCPG.

that the projection error for the mCPG algorithm is smaller than that for the CPG algorithm. This indicates that the cone constructed using the mCPG algorithm yields more accurate approximations than the one obtained by means of the CPG algorithm. To further compare the CPG and the mCPG algorithms, we introduce for all $r \geq 0$, the quantity $\widehat{e}_{\text{orth}}(r)$ which measures the orthogonality of the vector $\widehat{\nu}_{r+1}$, constructed by either the CPG or the mCPG algorithm, with respect to the cone $\widehat{K}_r$. This quantity is defined as

$$\widehat{e}_{\text{orth}}(r) := \|(\mathbb{I}^{\widehat{\mathcal{W}}} - \Pi^{\widehat{\mathcal{W}}}_{\widehat{K}_r})(\widehat{\nu}_{r+1})\|_{\widehat{\mathcal{W}}} \leq 1. \tag{41}$$

Notice that by definition, the larger $\widehat{e}_{\text{orth}}(r)$ (i.e., close to 1), the closer $\widehat{\nu}_{r+1}$ to being orthogonal to the cone. It can be seen in Figure 6 that $\widehat{e}^{\texttt{CPG}}_{\text{orth}}(r) \leq \widehat{e}^{\texttt{mCPG}}_{\text{orth}}(r)$ for all $r \geq 0$, which illustrates that the

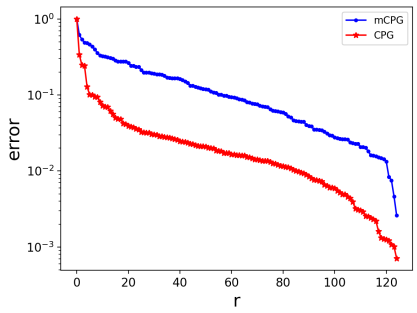


Figure 6: Obstacle test-case: Comparison between $\widehat{e}^{\texttt{CPG}}_{\text{orth}}(r)$ and $\widehat{e}^{\texttt{mCPG}}_{\text{orth}}(r)$.

basis constructed with the mCPG algorithm is of better quality than the one constructed with the CPG algorithm.

Considering the three tolerance pairs $(\delta_{\texttt{POD}} = 10^{-3}, \delta_{\texttt{mCPG}} = 10^{-2})$, $(\delta_{\texttt{POD}} = 10^{-4}, \delta_{\texttt{mCPG}} = 10^{-3})$, and $(\delta_{\texttt{POD}} = 10^{-6}, \delta_{\texttt{mCPG}} = 10^{-2})$, we obtain respectively the following pairs for the dimensions of the reduced bases: $(N, R) = (19, 50)$, $(N, R) = (51, 70)$, $(N, R) = (89, 50)$. With these choices, we cover the two possible settings, namely $N < R$ where we are sure that for all $\mu \in \mathcal{D}$, the bilinear form $\widehat{b}(\mu; \cdot, \cdot)$ is not inf-sup stable ($\widehat{\beta}^{\text{dec}}_{N,R}(\mu) = 0$) and $N \geq R$ where inf-sup stability cannot be asserted a priori. Our numerical results for the pair $(N, R) = (89, 50)$ show that the plain construction of the reduced model does not guarantee inf-sup stability in this case. Indeed, the value of the inf-sup constant $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ for the pair $(N, R) = (89, 50)$ is almost zero (of the order of 10^{-7}) for all $\mu \in \mathcal{D}_{\text{valid}}$.

Let us now consider the results obtained with the strategy proposed in Section 3. For all $\mu \in \mathcal{D}$,

we define

$$\widehat{\beta}_{\widehat{S}^n}(\mu) := \inf_{\widehat{\eta}\in\widehat{W}_R^+} \sup_{\widehat{v}\in\widehat{V}_N+\widehat{S}^n} \frac{b(\mu;\widehat{v},\widehat{\eta})}{\|\widehat{v}\|_{\widehat{\mathcal{V}}}\|\widehat{\eta}\|_{\widehat{\mathcal{W}}}}, \quad \forall n \in \mathbb{N}. \tag{42}$$

Figures 7c, 7f and 7i, respectively, show the values of $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$ as a function of $\mu \in \mathcal{D}_{\text{valid}}$ for the pairs $(N,R) = (19,50)$, $(N,R) = (51,70)$, $(N,R) = (89,50)$. For each pair, we consider the tight tolerance $\delta_{\texttt{PGA}} = 10^{-3}$ as well as the loose tolerance $\delta_{\texttt{PGA}} = 0.5$, 0.6, 0.95, respectively. These results illustrate the fact that the PGA algorithm does indeed recover the expected stability property. Notice that the values of $\mu \in \mathcal{D}_{\text{valid}}$ have been sorted in increasing order with respect to the inf-sup constant $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$. The panels in column 1 (resp. 2) of Figure 7 report the values of $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$ (see (23)) (resp. $\min_{\mu\in\mathcal{D}_{\texttt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$) as a function of n for the pairs $(N,R) = (19,50)$, $(N,R) = (51,70)$ and $(N,R) = (89,50)$ with $\delta_{\texttt{PGA}} = 10^{-3}$. The results show that the PGA algorithm converges and employs an effective greedy selection of supremizers in order to recover inf-sup stability for the reduced model. Figure 8 shows a comparison between the inf-sup constants $\widehat{\beta}_{\text{HF}}(\mu)$, $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$, $\widehat{\beta}^{\text{on}}_{N,R}(\mu)$, and $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$ for each of the reduced basis dimension pairs $(N,R) = (19,50)$, $(N,R) = (51,70)$ and $(N,R) = (89,50)$ with respectively $\delta_{\texttt{PGA}} = 0.5$, $\delta_{\texttt{PGA}} = 0.6$ and $\delta_{\texttt{PGA}} = 0.95$. On the one hand, we see that in all cases, in agreement with the theoretical results, $\widehat{\beta}^{\text{on}}_{N,R}(\mu) \geq \widehat{\beta}^{\text{off}}_{N,R}(\mu) \geq \widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ and $\widehat{\beta}^{\text{on}}_{N,R}(\mu) \geq \widehat{\beta}_{\text{HF}}(\mu)$ for all $\mu \in \mathcal{D}_{\text{valid}}$. On the other hand, we notice that there is no established order between $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$ and $\widehat{\beta}_{\text{HF}}(\mu)$.

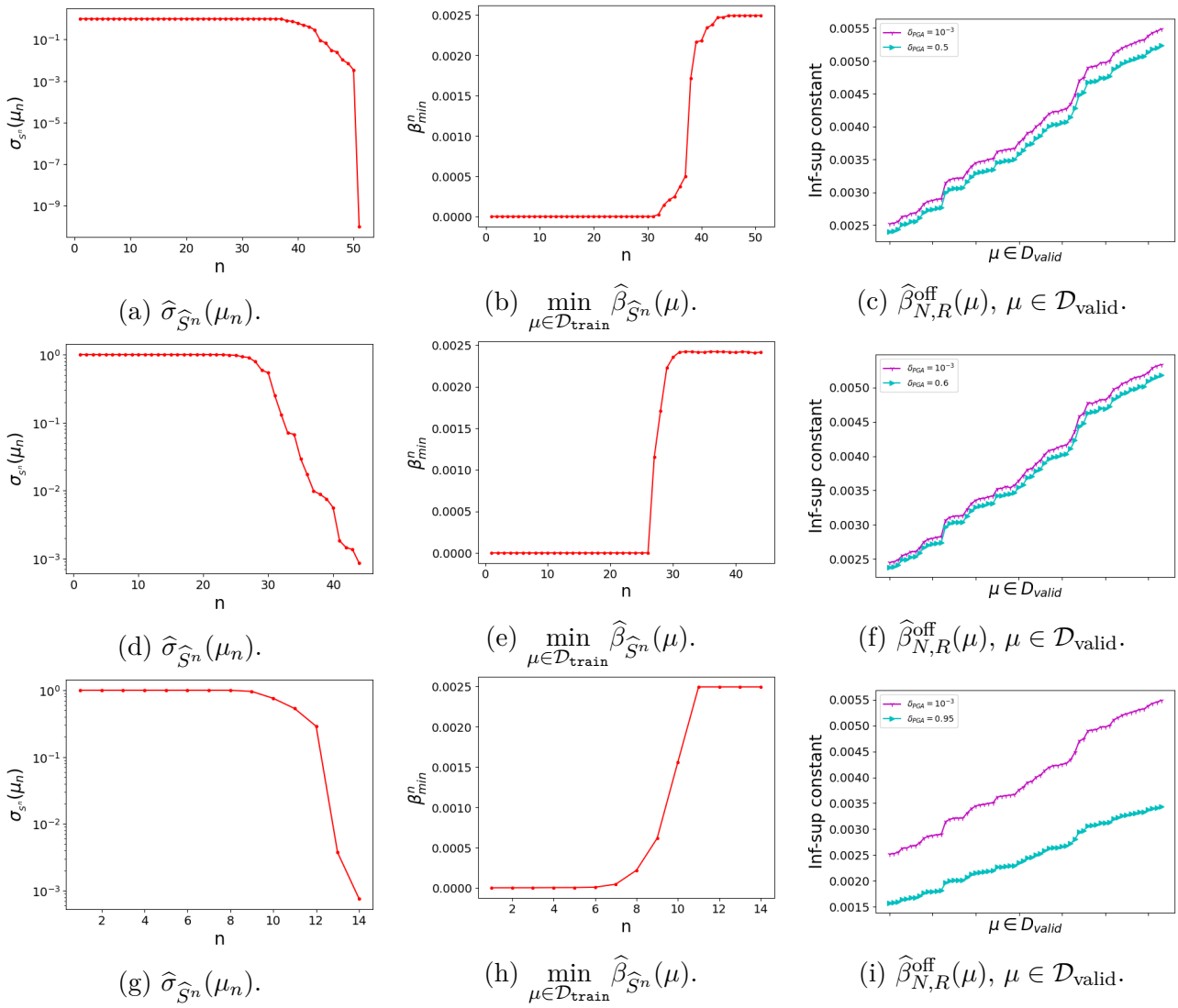
Table 1 shows that the offline phase of the offline enrichment method is approximately up to two times more expensive than the online enrichment method for the three reduced basis dimension pairs (N,R) with $\delta_{\texttt{PGA}} = 10^{-3}$. We see the opposite effect concerning the cost of the online phase. This is explained on the one hand by the fact that in the case of the online enrichment strategy, the primal basis is reconstructed for each calculation of the reduced model. On the other hand, it is also explained by the fact that in the present test-case, the dimension of the primal basis is smaller in the case of an offline enrichment strategy because the PGA algorithm converges in at most R iterations. We notice that the larger R , the more advantageous the offline strategy compared to the online strategy. It can also be seen that in the case of the offline strategy, the cost of the online phase is practically the same for all the three reduced basis dimension pairs, although the dimensions are different. This is explained by the fact that, once the reduced model is constructed, the cost of solving the resulting system varies quite moderately as a function of the dimension of the reduced bases. Comparing the overall cost in computation time between the two enrichment strategies, we conclude that the offline enrichment strategy is more advantageous than the online enrichment strategy if more than 102, 83, 52 online calculations are required with the reduced model for the reduced basis dimension pairs $(N,R) = (19,50)$, $(N,R) = (51,70)$, $(N,R) = (89,50)$, respectively. These numbers are called benefit threshold in Table 1.

5.2 Hertz contact between two half-disks

We consider two half-disks as shown in Figure 9. The first half-disk occupies a domain $\Omega_1 \subset \mathbb{R}^2$ of fixed radius $R_1 = 1\text{m}$, and the second a domain $\Omega_2(\mu) \subset \mathbb{R}^2$ of parametric radius $R_2 := \mu$ with

$$\mu \in \mathcal{D} := [0.7, 1.3](\text{m}). \tag{43}$$

For reasons of symmetry, only quarter-disks are discretized. The initial gap between the two disks is equal to $\gamma_0 > 0$. We impose a displacement of $-d$ (resp. d) on $\Gamma_1^{\texttt{top}}$ (resp. $\Gamma_2^{\texttt{bot}}$) of Ω_1 (resp. of $\Omega_2(\mu)$), with $d \geq \frac{1}{2}\gamma_0$. We set $\Omega(\mu) := \Omega_1 \cup \Omega_2(\mu)$ as well as $E := 15\text{Pa}$ (resp. $\nu := 0.35$) for the Young modulus (resp. the Poisson coefficient). For all $\mu \in \mathcal{D}$, we assume that there is a smooth invertible geometric mapping $h(\mu) : \widehat{\Omega} \to \Omega(\mu)$ defined on a reference domain $\widehat{\Omega} := \widehat{\Omega}_1 \cup \widehat{\Omega}_2$ such that $\{\widehat{\Omega}_1, \widehat{\Omega}_2\}$ is a partition of $\widehat{\Omega}$, $h(\mu)(\widehat{\Omega}_1) = \Omega_1$, $h(\mu)(\widehat{\Omega}_2) = \Omega_2(\mu)$ (notice that $h(\mu)|_{\widehat{\Omega}_1}$ is parameter-independent). We denote by



(a) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (b) $\min_{\mu \in \mathcal{D}_{\texttt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (c) $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\text{valid}}$.

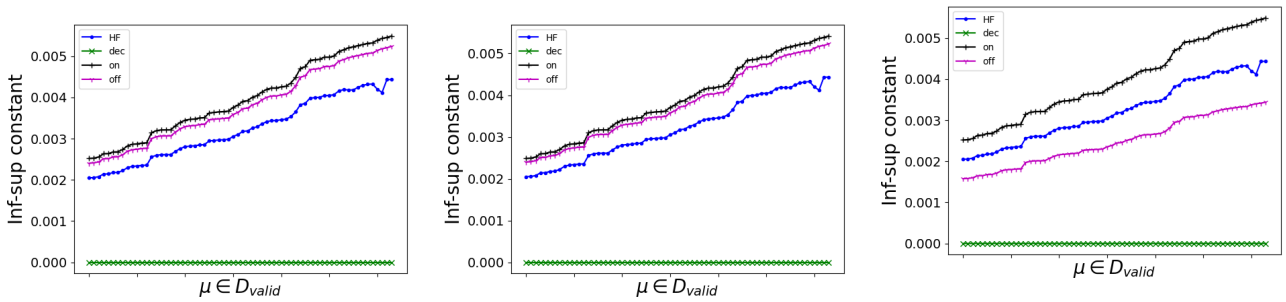
(d) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (e) $\min_{\mu \in \mathcal{D}_{\texttt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (f) $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\text{valid}}$.

(g) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (h) $\min_{\mu \in \mathcal{D}_{\texttt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (i) $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\text{valid}}$.

Figure 7: Obstacle test-case: PGA algorithm. Row 1: $(N, R) = (19, 50)$. Row 2: $(N, R) = (51, 70)$. Row 3 $(N, R) = (89, 50)$. Column 1 and Column 2: $\delta_{\texttt{PGA}} = 10^{-3}$. Column 3: $\delta_{\texttt{PGA}} \in \{0.95, 0.6, 0.5, 10^{-3}\}$.

(N, R)	$(19, 50)$		$(51, 70)$		$(89, 50)$	
Phase \ Method	on	off	on	off	on	off
Offline time(s)	1112	2296	1193	2395	1272	1884
Online time(s)	14	2.4	17	2.5	14	2.3
Benefit threshold	102		83		52	

Table 1: Obstacle test-case: The cost in seconds of the offline and online phases of the reduced model for the online (on) and offline (off) enrichment cases for the pairs $(N, R) = (19, 50)$, $(N, R) = (51, 70)$ and $(N, R) = (89, 50)$ with $\delta_{\texttt{PGA}} = 10^{-3}$.



(a) $(N, R) = (19, 50, \delta_{\texttt{PGA}} = 0.5)$. (b) $(N, R) = (51, 70, \delta_{\texttt{PGA}} = 0.6)$. (c) $(N, R) = (89, 50, \delta_{\texttt{PGA}} = 0.95)$.

Figure 8: Obstacle test-case: Inf-sup constant $\widehat{\beta}_{\mathrm{HF}}(\mu)$, $\widehat{\beta}^{\mathrm{dec}}_{N,R}(\mu)$, $\widehat{\beta}^{\mathrm{on}}_{N,R}(\mu)$, and $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$ for $\mu \in \mathcal{D}_{\mathrm{valid}}$.

$\widehat{\Gamma}^c_1$ (resp. $\widehat{\Gamma}^c_2$) the potential contact manifold located on $\partial\widehat{\Omega}_1$ (resp. $\partial\widehat{\Omega}_2$), and we set $\Gamma^c_1 := h(\mu)(\widehat{\Gamma}^c_1)$ and $\Gamma^c_2(\mu) := h(\mu)(\widehat{\Gamma}^c_2)$. The training set is chosen as $\mathcal{D}_{\mathrm{train}} := \{0.7 + 0.0075i, 0 \leq i \leq 80\}(\mathrm{m})$ (altogether

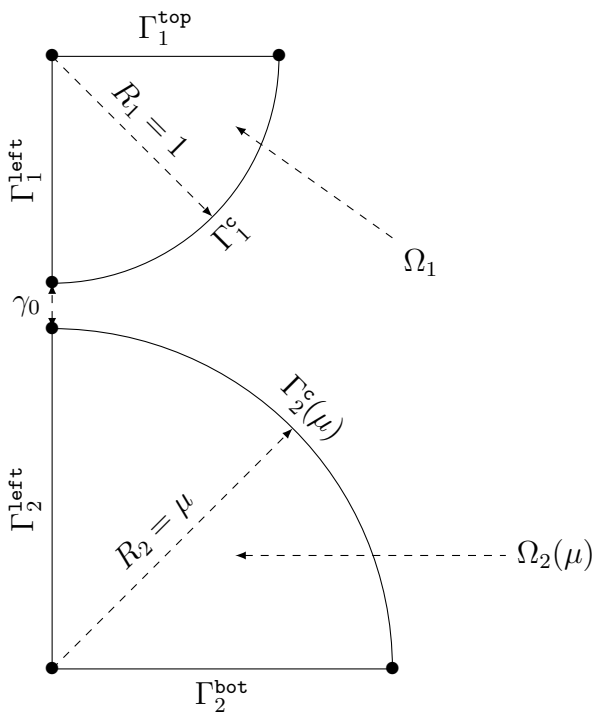


Figure 9: Hertz test-case: Configuration.

81 points) and the validation set $\mathcal{D}_{\mathrm{valid}}$ is generated by choosing 30 elements in $\mathcal{D}$ randomly with a uniform distribution.

We consider the HF subspace $\widehat{\mathcal{V}} \subset H^1(\widehat{\Omega}; \mathbb{R}^2)$ and the HF subcone $\widehat{\mathcal{W}}^+ \subset L^2(\widehat{\Gamma}^c_1, \mathbb{R}_+)$ built using $\mathbb{P}_1$ finite elements for displacement and the LAC (Local Average Contact) method [1] with $\mathbb{P}_0$ finite elements for the Lagrange multiplier on a mesh composed of 6130 nodes with 560 nodes on the potential contact manifold $\widehat{\Gamma}^c_1$. Thus, we have $\mathcal{N} := \dim(\widehat{\mathcal{V}}) = 12260$ primal and $\mathcal{R} := \dim(\widehat{\mathcal{W}}) = 280$ dual degrees of freedom. We equip the space $\widehat{\mathcal{V}}$ with the norm $\|\cdot\|_{\widehat{\mathcal{V}}}$ defined as follows (we use boldface

notation for vector-valued fields):

$$\|\widehat{\boldsymbol{v}}\|_{\widehat{\mathcal{V}}} := \left(\|\widehat{\boldsymbol{v}}\|_{L^2}^2 + \hat{\ell}^2\|\nabla\widehat{\boldsymbol{v}}\|_{L^2}^2\right)^{\frac{1}{2}}, \quad \forall\widehat{\boldsymbol{v}} \in \widehat{\mathcal{V}}, \tag{44}$$

where $\hat{\ell}$ is a characteristic length of $\widehat{\Omega}$ which is introduced for dimensional consistency. The variational formulation of the contact problem leads to a minimization problem of the form (1) subject to a so-called non-interpenetration condition which can be written either under the small deformation assumption (linear) or under the large deformation assumption (nonlinear). The bilinear form $\widehat{a}(\mu;\cdot,\cdot) : \widehat{\mathcal{V}} \times \widehat{\mathcal{V}} \to \mathbb{R}$ associated with the equilibrium equation in $\Omega(\mu)$ is defined by

$$\widehat{a}(\mu;\widehat{\boldsymbol{u}},\widehat{\boldsymbol{v}}) := \int_{\Omega(\mu)} \sigma(\varepsilon(\widehat{\boldsymbol{u}} \circ h(\mu)^{-1})) : \varepsilon(\widehat{\boldsymbol{v}} \circ h(\mu)^{-1})\, d\Omega(\mu), \tag{45}$$

where the linearized strain tensor $\varepsilon(\widehat{\boldsymbol{v}})$ is given by

$$\varepsilon(\widehat{\boldsymbol{v}}) := \frac{1}{2}(\nabla\widehat{\boldsymbol{v}} + \nabla\widehat{\boldsymbol{v}}^\top), \tag{46}$$

and the stress tensor $\sigma(\varepsilon)$ is given by

$$\sigma(\varepsilon) := \frac{E\nu}{(1+\nu)(1-2\nu)} tr(\varepsilon)\mathbb{I}_2 + \frac{E}{(1+\nu)}\varepsilon, \tag{47}$$

with $\mathbb{I}_2$ the identity matrix of order 2. The linear form $\widehat{f}(\mu;\cdot) : \widehat{\mathcal{V}} \to \mathbb{R}$ associated with the external load $\boldsymbol{\ell}(\mu) : \Omega(\mu) \to \mathbb{R}^2$ applied to $\Omega(\mu)$ is defined as

$$\widehat{f}(\mu;\widehat{\boldsymbol{v}}) := \int_{\Omega(\mu)} \boldsymbol{\ell}(\mu) \cdot \left(\widehat{\boldsymbol{v}} \circ h(\mu)^{-1}\right) d\Omega(\mu). \tag{48}$$

We consider here the first situation where we make the assumption of small deformations so that $\widehat{\Omega}$ deviates very little from its initial configuration. This allows us to assume that the normal vector is constant on the whole contact manifold. We denote by $\boldsymbol{n} := \boldsymbol{e}_y = (0,1)^\top$ the outward normal vector in $\widehat{\Gamma}_1^c$ and by $\widehat{\boldsymbol{\phi}} : \widehat{\Gamma}_1^c \to \widehat{\Gamma}_2^c$ the pairing function. The non-interpenetration condition is written as follows:

$$(\widehat{\boldsymbol{u}}_1(\mu)(\widehat{\boldsymbol{x}}) - \widehat{\boldsymbol{u}}_2(\mu) \circ \widehat{\boldsymbol{\phi}}(\widehat{\boldsymbol{x}})) \cdot \boldsymbol{e}_y \leq (\widehat{\boldsymbol{\phi}}(\widehat{\boldsymbol{x}}) - \widehat{\boldsymbol{x}}) \cdot \boldsymbol{e}_y, \quad \forall\widehat{\boldsymbol{x}} \in \widehat{\Gamma}_1^c. \tag{49}$$

Thus, the admissible set $\widehat{\mathcal{K}}(\mu)$ becomes

$$\widehat{\mathcal{K}}(\mu) := \{\widehat{\boldsymbol{v}} \in \widehat{\mathcal{V}} \mid \widehat{b}(\mu;\widehat{\boldsymbol{v}},\widehat{\eta}) \leq \widehat{g}(\mu;\widehat{\eta}),\ \forall\widehat{\eta} \in \widehat{\mathcal{W}}^+\}, \tag{50}$$

where

$$\widehat{b}(\mu;\widehat{\boldsymbol{v}},\widehat{\eta}) := \int_{\widehat{\Gamma}_1^c} (\widehat{\boldsymbol{v}}_1(\mu)(\widehat{\boldsymbol{x}}) - (\widehat{\boldsymbol{v}}_2(\mu) \circ \widehat{\boldsymbol{\phi}})(\widehat{\boldsymbol{x}})) \cdot \boldsymbol{e}_y\widehat{\eta}(\widehat{\boldsymbol{x}})\, d\widehat{\boldsymbol{x}}, \tag{51}$$

$$\widehat{g}(\mu;\widehat{\eta}) = \widehat{g}(\widehat{\eta}) := \int_{\widehat{\Gamma}_1^c} (\widehat{\boldsymbol{\phi}}(\widehat{\boldsymbol{x}}) - \widehat{\boldsymbol{x}}) \cdot \boldsymbol{e}_y\widehat{\eta}(\widehat{\boldsymbol{x}})\, d\widehat{\boldsymbol{x}}. \tag{52}$$

The initial gap γ_0 and the imposed displacement d are, respectively, set to $\gamma_0 := 0.001$m and $d := 0.09$m. This latter value, which is less than 10% of the maximum value between R_1 and R_2, allows us to remain within the validity of the small deformation assumption.

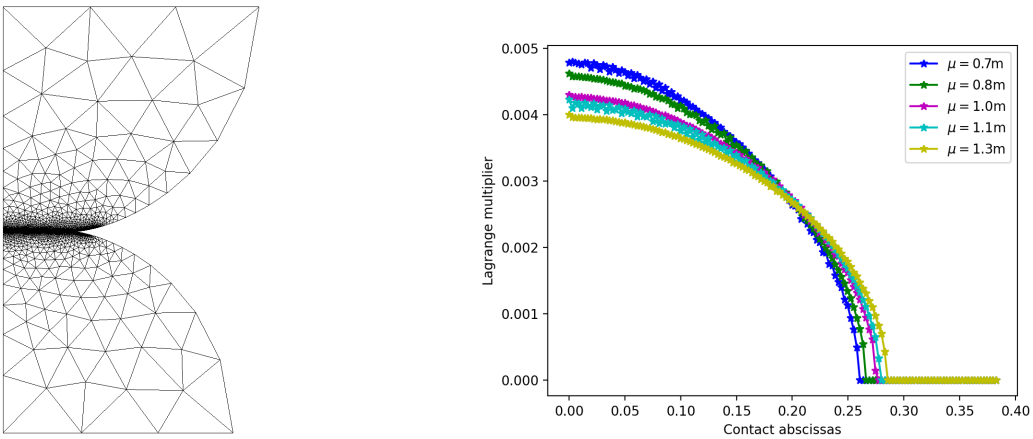


Figure 10: Hertz test-case: HF solution. Left: $\boldsymbol{u}(\mu)$ for $\mu = 0.9m$. Right: $\lambda(\mu)$ for $\mu \in \Big\{0.7, 0.8, 1.0, 1.1, 1.3\Big\}(m)$.

Figure 10 displays the deformed configuration resulting from the HF displacement field $\boldsymbol{u}(\mu) := \widehat{\boldsymbol{u}}(\mu) \circ h(\mu)^{-1}$ for $\mu = 0.9$m on its left panel and the Lagrange multiplier $\lambda(\mu) := \widehat{\lambda}(\mu) \circ h(\mu)^{-1}$ as a function of the abscissa along $\widehat{\Gamma}_1^c$ for $\mu \in \{0.7, 0.8, 1.0, 1.1, 1.3\}(m)$ on its right panel. We notice that the non-interpenetration condition is respected on the deformed configuration as expected. We can also notice some small oscillations in the Lagrange multipliers. This is due to the fact that for the evaluation of the contact contributions, a direct projection of the integration scheme is used instead of a projection-intersection of the master cells onto the slave cells. Figure 11 shows on its left (resp. right) panel the projection error $\widehat{e}_{\texttt{POD}}$ (resp. $\widehat{e}_{\texttt{CPG}}$ and $\widehat{e}_{\texttt{mCPG}}$) defined in (11) (resp. (13) and (34)) produced by the POD (resp. CPG and mCPG) algorithms as a function of the number of vectors composing the primal (resp. dual) reduced bases. In all cases, the projection errors decrease sufficiently fast so that the linear spaces generated by the primal and dual snapshots can be approximated by small-dimensional subspaces. We also notice that the projection error for the mCPG algorithm is smaller

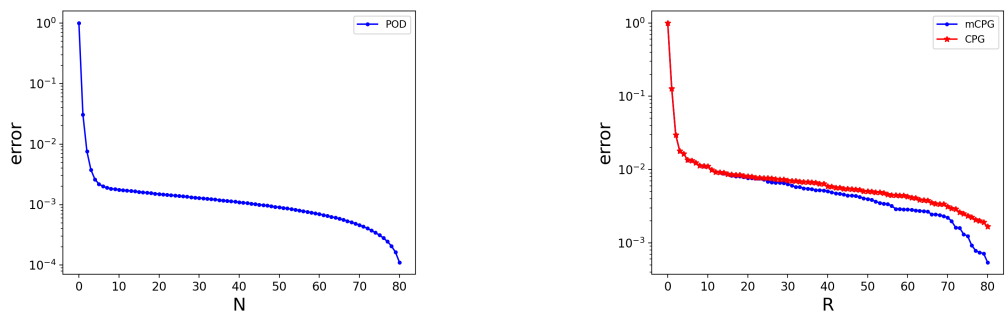


Figure 11: Hertz test-case: projection errors for the POD, CPG and mCPG algorithms. Left: POD. Right: CPG and mCPG.

than that for the CPG algorithm. This numerically indicates that the cone constructed using the mCPG algorithm yields more accurate approximations than the one obtained by means of the CPG algorithm. As before, it can be seen in Figure 12 that $\widehat{e}_{\text{orth}}^{\texttt{CPG}}(r) \leq \widehat{e}_{\text{orth}}^{\texttt{mCPG}}(r)$ for all $r \geq 0$, which illustrates that the basis constructed with the mCPG algorithm is of better quality than the one constructed with the CPG algorithm.

Considering the three tolerance pairs $(\delta_{\texttt{POD}} = 10^{-3}, \delta_{\texttt{mCPG}} = 10^{-2})$, $(\delta_{\texttt{POD}} = 10^{-3}, \delta_{\texttt{mCPG}} = 10^{-3})$,

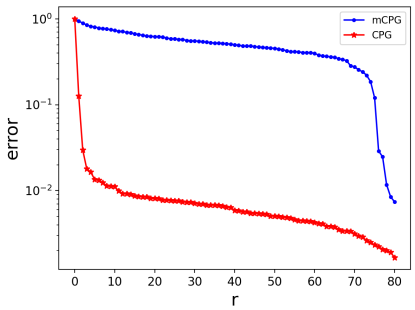


Figure 12: Hertz test-case: Comparison between $\widehat{e}^{\texttt{CPG}}_{\text{orth}}(r)$ and $\widehat{e}^{\texttt{mCPG}}_{\text{orth}}(r)$.

and $(\delta_{\texttt{POD}} = 10^{-2}, \delta_{\texttt{mCPG}} = 10^{-2})$, we obtain respectively the following pairs for the dimensions of the reduced bases: $(N = 46, R = 12)$, $(N = 46, R = 70)$, $(N = 3, R = 12)$. With these choices, we cover the two possible settings, namely $N < R$ where we are sure that for all $\mu \in \mathcal{D}$, the bilinear form $\widehat{b}(\mu; \cdot, \cdot)$ is not inf-sup stable ($\widehat{\beta}^{\text{dec}}_{N,R}(\mu) = 0$) and $N \geq R$ where inf-sup stability cannot be asserted a priori. Figure 13 shows the coefficient $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ as a function of $\mu \in \mathcal{D}_{\text{valid}}$ for the pair $(N, R) = (46, 12)$. Contrary to the results obtained in the obstacle test-case, the value of the inf-sup constant $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ is non-zero although it remains very small compared with $\widehat{\beta}_{\text{HF}}(\mu)$ (which is of the order of 8).

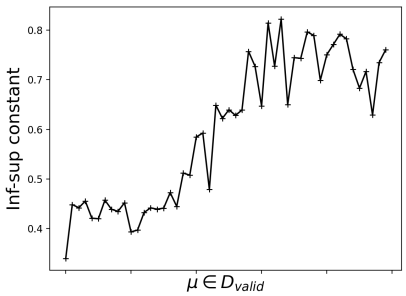
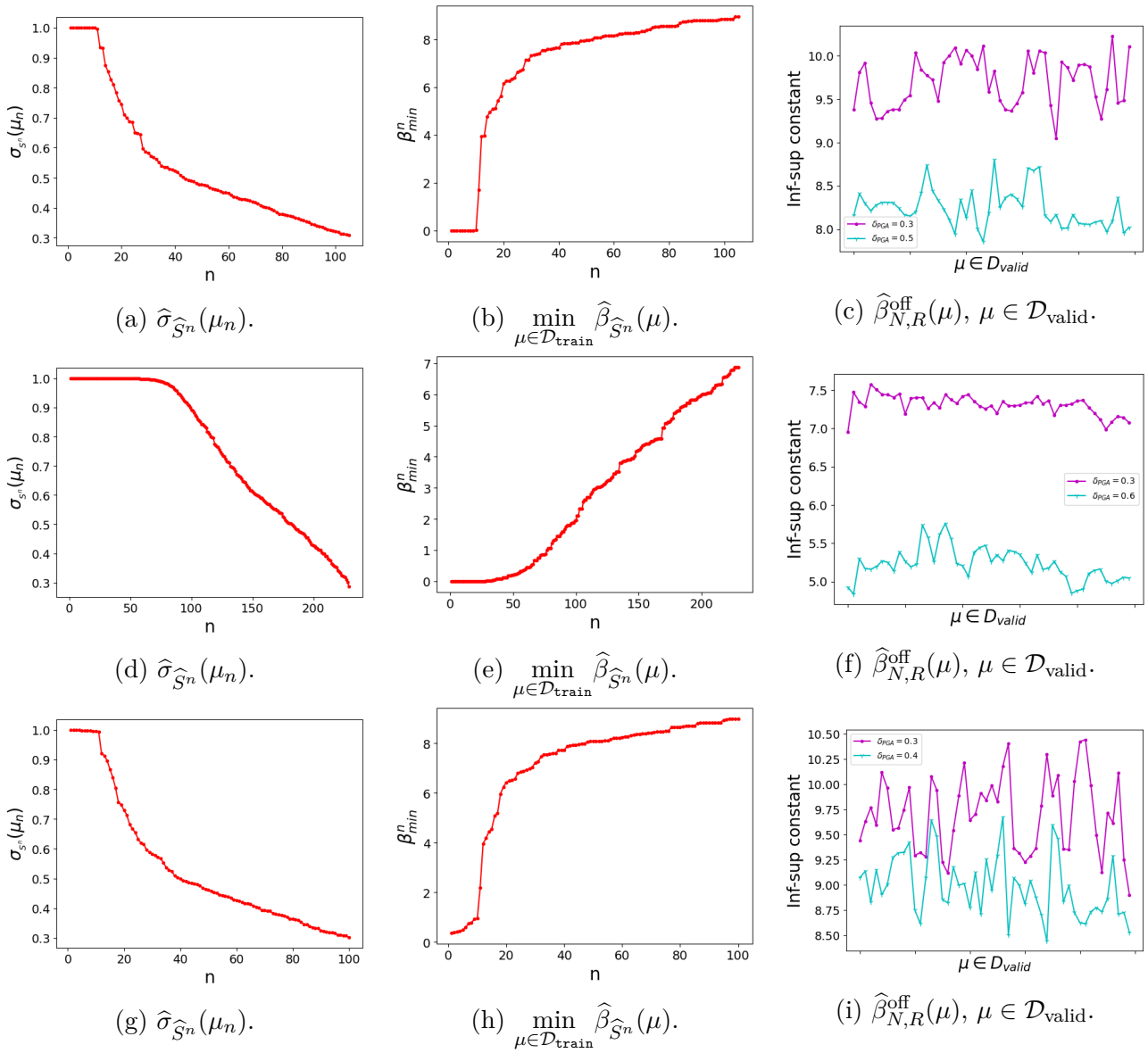


Figure 13: Hertz test-case: Inf-sup constant $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ with $(N, R) = (46, 12)$.

Let us now consider the results obtained with the strategy proposed in Section 3. Figures 14c, 14f and 14i, respectively, show the values of $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$ as a function of $\mu \in \mathcal{D}_{\text{valid}}$ for the pairs $(N, R) = (3, 12)$, $(N, R) = (46, 70)$, $(N, R) = (46, 12)$. For each of the above pairs, we consider the tighter tolerance $\delta_{\texttt{PGA}} = 0.3$ as well as the looser tolerance $\delta_{\texttt{PGA}} = 0.5$, 0.6, 0.4, respectively. These results illustrate the fact that the PGA algorithm does indeed recover the expected stability property. The panels in column 1 (resp. 2) of Figure 14, report the values of $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$ (see (23)) (resp. $\min_{\mu \in \mathcal{D}_{\texttt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$ (see (42))) as a function of n for the pairs $(N, R) = (3, 12)$, $(N, R) = (46, 70)$, $(N, R) = (46, 12)$ with $\delta_{\texttt{PGA}} = 0.3$. The results show that the PGA algorithm converges and employs an effective greedy selection of supremizers in order to recover inf-sup stability for the reduced model. Another interesting observation is that the decrease of $\widehat{\sigma}_{\widehat{S}^n}(\mu^n)$ with respect to n is much slower than for the previous test-case.

Figure 15 shows a comparison between the inf-sup constants $\widehat{\beta}_{\text{HF}}(\mu)$, $\widehat{\beta}^{\text{dec}}_{N,R}(\mu)$, $\widehat{\beta}^{\text{on}}_{N,R}(\mu)$, and $\widehat{\beta}^{\text{off}}_{N,R}(\mu)$ for each of reduced basis dimension pairs $(N = 46, R = 12)$, $(N = 46, R = 70)$, and $(N = 3, R = 12)$ with respectively $\delta_{\texttt{PGA}} = 0.5$, $\delta_{\texttt{PGA}} = 0.6$ and $\delta_{\texttt{PGA}} = 0.4$. On the one hand, we see that in all cases, in agreement with the theoretical results, $\widehat{\beta}^{\text{on}}_{N,R}(\mu) \geq \widehat{\beta}^{\text{off}}_{N,R}(\mu) \geq \widehat{\beta}^{\text{dec}}_{N,R}(\mu)$ and



(a) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (b) $\min_{\mu \in \mathcal{D}_{\mathtt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (c) $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\mathrm{valid}}$.

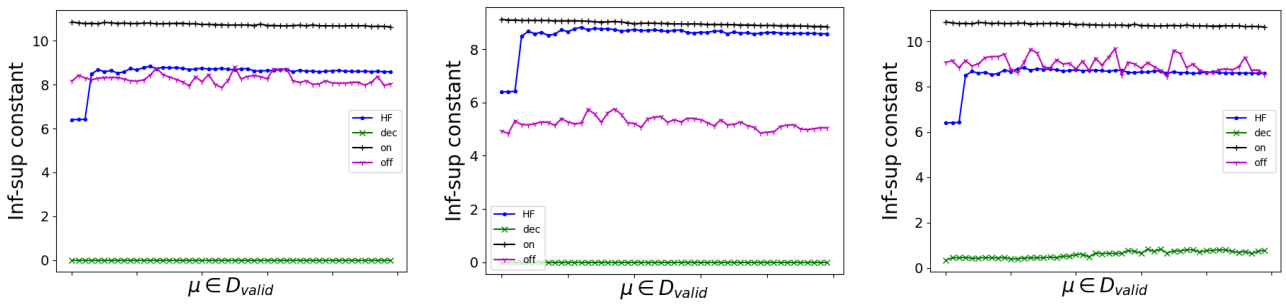
(d) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (e) $\min_{\mu \in \mathcal{D}_{\mathtt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (f) $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\mathrm{valid}}$.

(g) $\widehat{\sigma}_{\widehat{S}^n}(\mu_n)$. (h) $\min_{\mu \in \mathcal{D}_{\mathtt{train}}} \widehat{\beta}_{\widehat{S}^n}(\mu)$. (i) $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$, $\mu \in \mathcal{D}_{\mathrm{valid}}$.

Figure 14: Hertz test-case: PGA algorithm. Row 1: $(N, R) = (3, 12)$. Row 2: $(N, R) = (46, 70)$. Row 3 $(N, R) = (46, 12)$. Column 1 and Column 2: $\delta_{\mathtt{PGA}} = 0.3$. Column 3: $\delta_{\mathtt{PGA}} \in \{0.6, 0.5, 0.4, 0.3\}$.

$\widehat{\beta}^{\mathrm{on}}_{N,R}(\mu) \geq \widehat{\beta}_{\mathrm{HF}}(\mu)$ for all $\mu \in \mathcal{D}_{\mathrm{valid}}$. On the other hand, we notice that there is no established order between $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$ and $\widehat{\beta}_{\mathrm{HF}}(\mu)$.

Finally, Table 2 shows that the offline phase of the offline enrichment method is up to two times more expensive than the online enrichment method for the three reduced basis dimension pairs (N, R) with $\delta_{\mathtt{PGA}} = 0.3$. Overall, the same conclusions can be reached as those related to the previous test-case in Table 1.



(a) $(N, R) = (3, 12), \delta_{\texttt{PGA}} = 0.5$. (b) $(N, R) = (46, 70), \delta_{\texttt{PGA}} = 0.6$. (c) $(N, R) = (46, 12), \delta_{\texttt{PGA}} = 0.4$.

Figure 15: Hertz test-case: Inf-sup constant $\widehat{\beta}_{\mathrm{HF}}(\mu)$, $\widehat{\beta}^{\mathrm{dec}}_{N,R}(\mu)$, $\widehat{\beta}^{\mathrm{on}}_{N,R}(\mu)$, and $\widehat{\beta}^{\mathrm{off}}_{N,R}(\mu)$ for $\mu \in \mathcal{D}_{\mathrm{valid}}$.

(N, R)	$(3, 12)$		$(46, 70)$		$(46, 12)$	
Phase \ Method	on	off	on	off	on	off
Offline time(s)	8490	18098	10889	19750	10874	18312
Online time(s)	75	18	162	20	78	19
Benefit threshold	169		62		126	

Table 2: Hertz test-case: The cost in seconds of the offline and online phases of the reduced model for the online (on) and offline (off) enrichment cases for the pairs $(N, R) = (3, 12)$, $(N, R) = (46, 70)$ and $(N, R) = (46, 12)$ with $\delta_{\texttt{PGA}} = 0.3$.

References

[1] M. Abbas, G. Drouet, and P. Hild. The local average contact (LAC) method. *Comput. Methods Appl. Mech. Engrg.*, 339:488–513, 2018.

[2] S. Ali, F. Ballarin, and G. Rozza. Stabilized reduced basis methods for parametrized steady Stokes and Navier-Stokes equations. *Comput. Math. Appl.*, 80(11):2399–2416, 2020.

[3] D. Amsallem and C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM J. Sci. Comput.*, 33(5):2169–2198, 2011.

[4] M. S. Andersen, J. Dahl, and L. Vandenberghe. Cvxopt: A python package for convex optimization. Open source on `http://www.abel.ee.ucla.edu/cvxopt`, 2008.

[5] E. Bader, Z. Zhang, and K. Veroy. An empirical interpolation approach to reduced basis approximations for variational inequalities. *Math. Comput. Model. Dyn. Syst.*, 22(4):345–361, 2016.

[6] A. Bakhta, T. Boiveau, Y. Maday, and O. Mula. Epidemiological forecasting with model reduction of compartmental models. Application to the covid-19 pandemic. *Biology*, 10(1):22, 2020.

[7] M. Balajewicz, D. Amsallem, and C. Farhat. Projection-based model reduction for contact problems. *Internat. J. Numer. Methods Engrg.*, 106(8):644–663, 2016.

[8] F. Ballarin, A. Manzoni, A. Quarteroni, and G. Rozza. Supremizer stabilization of POD-Galerkin approximation of parametrized steady incompressible Navier-Stokes equations. *Internat. J. Numer. Methods Engrg.*, 102(5):1136–1161, 2015.

[9] A. Benaceur, A. Ern, and V. Ehrlacher. A reduced basis method for parametrized variational inequalities applied to contact mechanics. *Internat. J. Numer. Methods Engrg.*, 121(6):1170–1197, 2020.

[10] F. Brezzi, W. W. Hager, and P.-A. Raviart. Error estimates for the finite element solution of variational inequalities. II. Mixed methods. *Numer. Math.*, 31(1):1–16, 1978/79.

[11] O. Burkovska, B. Haasdonk, J. Salomon, and B. Wohlmuth. Reduced basis methods for pricing options with the Black-Scholes and Heston models. *SIAM J. Financial Math.*, 6(1):685–712, 2015.

[12] K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations. *Internat. J. Numer. Methods Engrg.*, 86(2):155–181, 2011.

[13] A. Chatterjee. An introduction to the proper orthogonal decomposition. *Current Science*, pp. 808–817, 2000.

[14] W. Dahmen, C. Plesken, and G. Welper. Double greedy algorithms: Reduced basis methods for transport dominated problems. *ESAIM Math. Model. Numer. Anal.*, 48(3):623–663, 2014.

[15] R. S. Falk. Error estimates for the approximation of a class of variational inequalities. *Math. Comput.*, 28:963–971, 1974.

[16] J. Fauque, I. Ramière, and D. Ryckelynck. Hybrid hyper-reduced modeling for contact mechanics problems. *Internat. J. Numer. Methods Engrg.*, 115(1):117–139, 2018.

[17] G. Fichera. Elastostatics problems with unilateral constraints: The Signorini problem with ambiguous boundary conditions. pp. 613–679. Ediz. Cremonese, Rome, 1965.

[18] Gaetano Fichera. The Italian contribution to the mathematical theory of elasticity. *Meccanica–J. Italian Assoc. Theoret. Appl. Mech.*, 19(4):259–268, 1984.

[19] A.-L. Gerner and K. Veroy. Certified reduced basis methods for parametrized saddle point problems. *SIAM J. Sci. Comput.*, 34(5):A2812–A2836, 2012.

[20] S. Glas and K. Urban. Numerical investigations of an error bound for reduced basis approximations of noncoercice variational inequalities. *IFAC-PapersOnLine*, 48:721–726, 2015.

[21] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems.* Scientific Computation. Springer-Verlag, Berlin, 2008.

[22] B. Haasdonk, J. Salomon, and B. Wohlmuth. A reduced basis method for parametrized variational inequalities. *SIAM J. Numer. Anal.*, 50(5):2656–2676, 2012.

[23] F. Hecht. New development in freefem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[24] Heinrich Hertz. Ueber die Berührung fester elastischer Körper. *J. Reine Angew. Math.*, 92:156–171, 1882.

[25] Jan S. Hesthaven, Gianluigi Rozza, and Benjamin Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations.* SpringerBriefs in Mathematics. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2016.

[26] David Kinderlehrer and Guido Stampacchia. *An Introduction to Variational Inequalities and their Applications*, volume 31 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[27] A. S. Kravchuk and P. J. Neittaanmäki. *Variational and Quasi-Variational Inequalities in Mechanics*, volume 147 of *Solid Mechanics and its Applications*. Springer, Dordrecht, 2007.

[28] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, 90(1):117–148, 2001.

[29] D. D. Lee and H. S. Seung. Algorithms for nn-negative matrix factorization. NIPS'00, pp. 535–541. MIT Press, Cambridge, MA, USA, 2000.

[30] Y. Maday, A. T. Patera, and G. Turinici. A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. volume 17, pp. 437–446, 2002.

[31] R. Monneau. A brief overview on the obstacle problem. volume 202 of *Progr. Math.*, pp. 303–312. Birkhäuser, Basel, 2001.

[32] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, 26(1):17–32, 1981.

[33] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *J. Fluids Eng.* , 124(1):70–80, 2001.

[34] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations*, volume 92 of *Unitext*. Springer, Cham, 2016.

[35] A. Ribes and C. Caremoli. Salomé platform component model for numerical simulation. volume 2, pp. 553–564, 2007.

[36] G. Rozza, D. B. P. Huynh, and A. Manzoni. Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. *Numer. Math.*, 125(1):115–152, 2013.

[37] G. Rozza and K. Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *Comput. Methods Appl. Mech. Engrg.*, 196(7):1244–1260, 2007.

[38] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, 40(11):2323–2330, 2002.

[39] P. Wriggers. Finite element algorithms for contact problems. *Arch. Comput. Methods Engrg.*, 2(4):1–49, 1995.

[40] Z. Zhang, E. Bader, and K. Veroy. A slack approach to reduced-basis approximation and error estimation for variational inequalities. *C. R. Math. Acad. Sci. Paris*, 354(3):283–289, 2016.